novozymes®
Rethink Tomorrow

Ref.: 2018-05620-01
ALST/PNI

Glucoamylase from *Aspergillus niger*

An application to amend the *Australia New Zealand Food Standards Code* with a glucoamylase preparation produced by a genetically modified strain of *Aspergillus niger*

Novozymes A/S
June 2018

Novozymes A/S
Krogshøjvej 36
DK-2880 Bagsværd
Denmark

Telephone:
Telefax:

E-mail:
Internet:
www.novozymes.com

CVR number:
10 00 71 27

Table of Contents

EXECUTIVE SUMMARY

The present application seeks to amend Schedule 18 - Processing Aids of the Australia New Zealand Food Standards Code (the Code) to approve a glucoamylase enzyme preparation produced by Novozymes A/S.

Proposed change to Australia New Zealand Food Standards Code – Schedule 18 – Processing aids

The table in S18—4, Permitted enzymes (section 1.3.3—6)—Enzymes of microbial origin, is proposed to be amended to include a genetically modified strain of *Aspergillus niger* as permitted source for glucoamylase.

The application is applied for assessment by the general procedure.

Description of enzyme preparation

The enzyme is a glucan 1,4-alpha-glucosidase (EC 3.2.1.3), commonly known as glucoamylase.

Glucoamylases catalyse the hydrolysis of 1,4-alpha and 1,6-alpha-D-glucosidic linkages in starch polysaccharides.

The enzyme is produced by submerged fermentation of an *Aspergillus niger* microorganism expressing a glucoamylase from *Talaromyces emersonii*.

The enzyme product is commercially available as liquid or granulate preparations complying with the JECFA recommended purity specifications for food-grade enzymes.

The producing microorganism, *Aspergillus niger*, is absent from the commercial enzyme product.

Use of the enzyme

The glucoamylase preparation is used as a processing aid in the starch and alcohol industry (beverage alcohol and brewing) for saccharification of liquified starch. Generally, glucoamylase degrades starch into D-glucose.

- During starch processing to produce syrups, the glucoamylase degrades polysaccharides into glucose. The use of the enzyme leads to an increased glucose yield.
- In beverage alcohol (distilling) processes the glucoamylase is used in order to degrade gelatinised starch and dextrins into glucose and other fermentable sugars.
- During brewing and other cereal based beverage processes the glucoamylase converts starch into fermentable sugars for the production of beer.
- In baking and other cereal based processes, the glucoamylase hydrolyses starch, from milled and damaged granules, into glucose which will be fermented by yeast.
- During processing of fruits and vegetables the glucoamylase degrades starch present in the fruits and vegetables in order to increase yield and to facilitate processing.

Benefits

The benefits of the action of the glucoamylase in starch processing are:

- Efficient degradation of dextrins and production of glucose
- Reduced risk of contamination, because the enzyme can be used at high operating temperature
- Stable process allowing for variations in temperature

The benefits of the action of the glucoamylase in beverage alcohol (distilling) processes:

- Efficient degradation of dextrins and production of fermentable sugars
- High alcohol yields due to a more complete conversion of starch and thereby less use of raw materials
- Reduced risk of contamination, because the enzyme can be used at high operating temperature and low operating pH

The benefits of the action of the glucoamylase in brewing processes and other cereal based beverage processes:

- More uniform and predictable production process and brewing yield including the possibility to control the desired level of fermentable sugars at every production

The benefits of the action of the glucoamylase in baking processes and other cereal based processes:

- Reduced baking time and yeast boosting
- Uniform and slightly increased volume and enhanced crust colour of the bakery product
- More uniform and predictable production of glucose
- Improved fermentability
- Uniform colour

The benefits of the action of the glucoamylase in fruit and vegetable processing:

- Increased yield due to improved press capacity and filtration rates
- Avoidance of starch-based haziness in the final product giving a clearer product

Safety evaluation

The safety of the strain and the enzyme product has been thoroughly assessed:

- The production organism has a long history of safe use as production strain for food grade enzyme preparations and is known not to produce any toxic metabolites.
- The genetic modifications in the production strain are well-characterised and safe and the recombinant DNA is stably integrated into the production organism and unlikely to pose a safety concern.
- The enzyme preparation complies with international specifications ensuring absence of contamination by toxic substances or noxious microorganisms.
- Sequence homology assessment to known allergens and toxins shows that oral intake of the glucoamylase does not pose food allergenic or toxic concern.
- Two mutagenicity studies *in vitro* showed no evidence of genotoxic potential of the enzyme preparation.
- An oral feeding study in rats for 13-weeks showed that all dose levels were generally well tolerated and no evidence of toxicity.

Furthermore, the safety of the glucoamylase preparation was confirmed by external expert groups, as follows:

- Denmark: The enzyme preparation was safety assessed resulting in the authorisation of the enzyme product by the Danish Veterinary and Food Administration.
- Brazil: The enzyme was evaluated, approved and included in the Brazilian positive list – RDC 26/2009.
- Canada: The enzyme was evaluated by the Canadian authorities, Health Canada, and added to the Canada Food and Drug regulations, Division 16 Food Additives, Section B. 16.100, Table V, Food Additives that may be used as Food enzymes.
- China: The enzyme is included in the Chinese positive list – GB2760-2015.
- France: The enzyme is included in the French positive list for processing aids, including food enzymes (The French order of October 19, 2006 on use of processing aids in the manufacture of certain foodstuff), as amended.
- Mexico: Based on a dossier submitted by Novozymes, the Mexican food authorities, COFEPRIS, have approved the enzyme.

Conclusion

Based on the Novozymes safety evaluation (confirmed by the above-mentioned bodies), we respectfully request the inclusion of this enzyme in the table in S18—4, Permitted enzymes (section 1.3.3—6)—Enzymes of microbial origin.

INTRODUCTION

The present dossier describes a glucoamylase enzyme preparation produced by submerged fermentation of an *Aspergillus niger* microorganism expressing a glucoamylase from *Talaromyces emersonii*.

The enzyme is a glucan 1,4-alpha-glucosidase (EC 3.2.1.3), commonly known as glucoamylase. The enzyme catalyses the hydrolysis of 1,4-alpha and 1,6-alpha-D-glucosidic linkages in starch polysaccharides.

The glucoamylase enzyme preparation is intended to be used as a processing aid in the starch-processing industry to hydrolyse starch for the production of several products, e.g. syrups, distilled beverages, beer, bread, and juices.

The following sections describe in detail the construction of the genetically modified *Aspergillus niger* used as the production organism, the production process, the product specification, the application of the enzyme preparation and finally the safety evaluation of the product including the toxicology program, which has been carried out confirming the safety of the product for its intended use.

The documentation has been elaborated according to the Application Handbook from Food Standards Australia New Zealand as of March 1st 2016, applied as relevant for an enzyme application, i.e. outlining the following section:

- Section 3.1.1 – General requirements
- Section 3.3.2 – Processing aids, subsections A, C, D, E, F

NB! When reading this document it should be noticed that in some reports, the glucoamylase enzyme preparation is described by the internal production batch code PPQ36551 or NAFR 362.

CHAPTER 3.1, GENERAL REQUIREMENTS FOR APPLICATIONS

A. Executive Summary

An Executive Summary is provided as a separate copy together with this application.

B. Applicant details

(a) Applicant's name/s

(b) Company/organisation name
Novozymes Australia Pty Ltd

(c) Address (street and postal)
3/22 Loyalty Road PO Box 4942
2151 NORTH ROCKS NSW, Australia

(d) Telephone number

(e) Email address

(f) Nature of applicant's business
Biotechnology

(g) Details of other individuals, companies or organisations associated with the application.
Dossier prepared by:

Regulatory Affairs
Krogshoejvej 36
2880 Bagsvaerd Denmark
Mobile:
E-mail:

C. Purpose of the application

This application is submitted to provide for amendment of the Australia New Zealand Food Standards Code – Schedule 18 – Processing aids, Table S18—4 to include a genetically modified strain of *Aspergillus niger* as permitted source for a glucoamylase.

D. Justification for the application

The need for the proposed change

Table S18—4 of Schedule 18 contains a list of permitted enzymes of microbial origin. There are a number of approved glucoamylases EC 3.2.1.3 from different sources, including *Aspergillus niger*. However, Table S18—4 of schedule 18 does not contain a glucoamylase EC 3.2.1.3 from *Aspergillus niger* containing the gene for glucoamylase from *Talaromyces emersonii*.

Aspergillus niger is an approved host and production strain for a number of enzymes in Schedule 18, e.g. a wide range of enzymes that can be used in starch processing such as alpha-amylase, alpha-arabinofuranosidase, cellulase, endo-1,4-beta-xylanase, endo-arabinase, alpha-galactosidase, beta-galactosidase, beta-glucanase, glucoamylase, alpha-glucosidase, beta-glucosidase, hemicellulase multicomponent enzyme.

The advantages of the proposed change over the status quo

The glucoamylase preparation is used as a processing aid during the manufacture of starch-based products. Glucoamylases convert starch by removing D-glucose units in a stepwise manner from the non-reducing end of the substrate molecule to produce glucose for further processing to a wide range of products, such as syrup, distilled alcohol, beer, bread, and juices.

The benefits of the action of the glucoamylase in starch processing are:

- Efficient degradation of dextrins and production of glucose
- Reduced risk of contamination, because the enzyme can be used at high operating temperature
- Stable process allowing for variations in temperature

The benefits of the action of the glucoamylase in beverage alcohol (distilling) processes:

- Efficient degradation of dextrins and production of fermentable sugars
- High alcohol yields due to a more complete conversion of starch and thereby less use of raw materials
- Reduced risk of contamination, because the enzyme can be used at high operating temperature and low operating pH

The benefits of the action of the glucoamylase in brewing processes and other cereal based beverage processes:

- More uniform and predictable production process and brewing yield including the possibility to control the desired level of fermentable sugars at every production

The benefits of the action of the glucoamylase in baking processes and other cereal based processes:

- Reduced baking time and yeast boosting

- Uniform and slightly increased volume and enhanced crust colour of the bakery product
- More uniform and predictable production of glucose
- Improved fermentability
- Uniform colour

The benefits of the action of the glucoamylase in fruit and vegetable processing:

- Increased yield due to improved press capacity and filtration rates
- Avoidance of starch-based haziness in the final product giving a clearer product

The benefits, which are described above, are not exclusively obtainable by means of enzyme treatment but can be achieved without the use of enzymes, or with a reduced use of enzymes, through e.g. modified maybe more expensive or less environmentally friendly production processes or recipe changes.

As a response to international customer interests, registration activities have been done globally, e.g. the glucoamylase preparation has been approved in Denmark and evaluated and approved for the described applications in Brazil, Canada, China, France, and Mexico.

D.1 Regulatory impact information

D.1.1 Costs and benefits of the application

The application is not likely to place costs or regulatory restrictions on industry or consumers. Inclusion of the glucoamylase enzyme in Table S18—4 of Schedule 18 will provide the food and beverage industry with the opportunity to improve the yield of fermentable sugars for baking, distilling, and brewing. The glucoamylase will furthermore improve the yield and facilitate processing in the juice industry. The glucoamylase can achieve these benefits under environmentally friendly and cost efficient production conditions. For the government, the burden is limited to necessary activities for a variation of Schedule 18.

D.1.2 Impact on international trade

The application is not likely to cause impact on international trade.

E. Information to support the application

E.1 Data requirements

No public health and safety issues related to the proposed change are foreseen. As outlined in sections 3.3.2 C, D, E, F, the glucoamylase is produced by submerged fermentation of a genetically modified *Aspergillus niger* strain.

The safety of the strain and the enzyme product has been thoroughly assessed:

- The production organism has a long history of safe use as production strain for food grade enzyme preparations and is known not to produce any toxic metabolites.
- The genetic modifications in the production strain are well-characterised and safe and the recombinant DNA is stably integrated into the production organism and unlikely to pose a safety concern.

- The enzyme preparation complies with international specifications ensuring absence of contamination by toxic substances or noxious microorganisms
- Sequence homology assessment to known allergens and toxins shows that oral intake of the glucoamylase does not pose food allergenic or toxic concern.
- Two mutagenicity studies *in vitro* showed no evidence of genotoxic potential of the enzyme preparation.
- An oral feeding study in rats for 13-weeks showed that all dose levels were generally well tolerated and no evidence of toxicity.

F. Assessment procedure

Because the application is for a new source organism for an existing enzyme in the Code, it is considered appropriate that the assessment procedure is characterised as “General Procedure, Level 1”.

G. Confidential commercial information (CCI)

Detailed information on the construction and characteristics of the genetically modified production strain is provided in Appendix 6. A summary of this information is given in Section 3.3.2 E. The formal request for treatment of selected parts of Appendix 6 as confidential commercial information (CCI) is included as Appendix 1.1.

H. Other confidential information

Apart from the selected parts of Appendix 6 identified as confidential commercial information (CCI), no other information is requested to be treated as confidential.

I. Exclusive capturable commercial benefit (ECCB)

This application is not expected to confer an Exclusive Capturable Commercial Benefit.

J. International and other national standards

J.1 International Standards

Use of enzymes as processing aids for food production is not restricted by any Codex Alimentarius Commission (Codex) Standards.

J.2 Other national standards or regulations

With few exceptions on national, commodity standards, use of enzymes as processing aids for food production is in general not restricted by standards or regulations in other countries.

K. Statutory declaration

The Statutory Declaration is provided as a separate document together with this submission.

L. Checklist

This application concerns an enzyme product intended to be used as a processing aid. Therefore, the relevant documentation according to the Application Handbook from Food Standards Australia New Zealand as of March 1st 2016, are the following sections:

- Section 3.1.1 – General requirements
- Section 3.3.2 – Processing aids, subsections A, C, D, E, F

Accordingly, the checklist for General requirements as well as the Processing aids part of the checklist for applications for substances added to food was used and is included as Appendix 1.2 and 1.3.

CHAPTER 3.3, GUIDELINES FOR APPLICATIONS FOR SUBSTANCES ADDED TO FOOD

3.3.2 PROCESSING AIDS

The glucoamylase enzyme preparation described in this application is representative for the commercial food enzyme product for which approval is sought.

A. Technical information on the processing aid

A.1. Information on the type of processing aid

The glucoamylase enzyme preparation belongs to the category of processing aids described in Table S18—4 of Schedule 18, Enzymes of microbial origin.

The glucoamylase enzyme preparation is to be used in the food industry as a processing aid during the processing of raw materials containing starch. Glucoamylase converts starch to glucose.

The glucoamylase enzyme preparation is used in, but not limited to, the following food manufacturing processes:

- In starch processing to produce syrups, the glucoamylase degrades polysaccharides into glucose. The use of the enzyme leads to a higher glucose yield.
- In beverage alcohol (distilling) processes the glucoamylase is used in order to degrade gelatinised starch and dextrins into glucose and other fermentable sugars.
- During brewing and other cereal based beverage processes the glucoamylase converts starch into fermentable sugars for the production of beer.
- In baking and other cereal based processes, the glucoamylase hydrolyses starch, from milled and damaged granules, into glucose which will be fermented by yeast.
- During processing of fruits and vegetables the glucoamylase degrades starch present in the fruits and vegetables in order to increase yield and to facilitate processing.

The highest dosage of the glucoamylase during a food manufacturing process is in brewing and other cereal based beverage processes, where dosages up to 4100 AGU per kg starch dry matter are used.

Examples of benefits when applying the glucoamylase enzyme preparation during brewing are provided in the application information sheet (Appendix 2.1), in which the commercial name Attenuzyme Core is used to describe the glucoamylase enzyme preparation.

A.2. Information on the identity of the processing aid

A.2.1. Enzyme

Generic name:	Glucoamylase
IUB nomenclature:	Glucan 1,4-alpha-glucosidase
IUB No.:	EC 3.2.1.3
CAS No.:	9032-08-0

A.2.2. Enzyme preparation

The glucoamylase enzyme preparation is available as a single enzyme formulation under the commercial name Attenuzyme Core when used for brewing processes. Attenuzyme Core is used as an example for the typical glucoamylase enzyme preparation.

The Product Data Sheet for the commercial product Attenuzyme Core is enclosed as Appendix 2.2. The typical composition of Attenuzyme Core is shown below:

	Attenuzyme Core
Enzyme solids (TOS[a])	42.0 %
Sucrose/Glucose	1.0 %
Organic acid	0.0 %
Ash	0.0 %
Water	57.0 %

The commercial product Attenuzyme Core is standardised in glucoamylase units to an activity of 1600 AGU/g. The Novozymes method used to determine the AGU activity is enclosed in Appendix 3.1.

Briefly, glucoamylase converts maltose to D-glucose and the reaction is stopped with an alkaline solution. The glucose is subsequently phosphorylated and oxidised by other enzymes during which an amount of NAD^+ proportional to maltose is reduced to NADH. This reduction reaction can be quantified following the increase in absorbance at 340 nm. The increase is proportional to the enzyme activity.

A.2.3. Host organism

The host strain is a modified (protease deficient) *Aspergillus niger* strain (BO-1) derived from a natural isolate of *Aspergillus niger* C40-1. The BO-1 strain lineage has been used by Novozymes for more than 30 years and has given rise to a number of food enzyme production strains, which are used for production of previously evaluated and regulatory approved food enzymes. The taxonomic classification of the strain is as follows:

Name:	*Aspergillus niger*
Class:	Eurotiomycetes
Order:	Eurotiales
Genus:	*Aspergillus*
Species:	*niger*

For a more detailed description of the host organism and the genetic modifications, please see Section 3.3.2 E.

[a] TOS = Total Organic Solids, defined as: 100% - water - ash - diluents

A.2.4. Donor organism

The glucoamylase is from *Talaromyces emersonii.* The introduced gene was amplified by PCR from genomic DNA isolated from the donor organism and inserted into two standard expression vectors. The glucoamylase is not protein engineered.

For a more detailed description of the donor and the donor gene, please see Section 3.3.2 E.

A.3. Information on the chemical and physical properties of the processing aid

The enzyme is a glucan 1,4-alpha-glucosidase (EC 3.2.1.3), commonly known as glucoamylase. Glucoamylases catalyse the hydrolysis of 1,4-alpha and 1,6-alpha-D-glucosidic linkages in starch polysaccharides.

The enzyme preparation is available as a liquid or granulated product. The liquid product is stabilised with sucrose and glucose.

The food enzyme object of the present dossier is not added to final foodstuffs but used as a processing aid during food manufacturing.

No reaction products, which could not be considered normal constituents of the diet, are formed during the production or storage of the enzyme treated food.

A.4. Manufacturing process

The manufacturing process is composed of a fermentation process, a purification process, a formulation process and finally a quality control of the finished product, as outlined by Aunstrup et al. (1979)[1]. This section describes the processes used in manufacturing of the glucoamylase enzyme product.

The enzyme preparation is manufactured in accordance with current Good Manufacturing Practices (Appendix 4.1). The quality management system used in the manufacturing process complies with ISO 9001:2015 (Appendix 4.2).

The raw materials are Food Grade Quality and have been subjected to appropriate analysis to ensure their conformity with the specifications.

A.4.1. Fermentation

The glucoamylase is produced by submerged fed-batch pure culture fermentation of the genetically modified strain of *Aspergillus niger*, described in Section 3.3.2 E.

A.4.1.1. Raw materials for fermentation

The production strain is grown in a medium consisting of compounds providing an adequate supply of carbon and nitrogen plus minerals and vitamins necessary for growth. The choice of raw materials used in the fermentation process (the feed, the seed fermenter, the main fermenter and dosing) is listed below.

- Potable water
- Carbohydrates (e.g. corn starch, glucose syrup, sucrose)
- Vegetable protein (e.g. soy bean meal)
- Ammonia
- Salts (e.g. $MgSO_4$)
- Trace metals (e.g. $NiCl_2$, $MnSO_4$, $FeSO_4$, $CuSO_4$, $ZnSO_4$)
- Alkali and acid for pH adjustments (e.g. citric acid, sulfuric acid, ammonia, sodium hydroxide)
- Antifoaming agents (e.g. polypropylene glycol, polyalkylene glycol)

A.4.1.2. Hygienic precautions

All equipment is designed and constructed to prevent contamination by foreign microorganisms.

All valves and connections not in use for the fermentation are sealed by steam at more than 120 °C.

After sterilisation a positive pressure of more than 0.2 atmosphere is maintained in the fermentation tank.

The air used for aeration is sterilised by passing a sterile filter.

The inside of each fermentation tank is cleaned between fermentations by means of a high-pressure water jet and inspected after the cleaning procedures have been completed.

A.4.1.3. Preparation of the inoculum

The inoculum flask containing the prepared medium is autoclaved and checked. Only approved flasks are used for inoculation.

The stock culture suspension is injected aseptically into the inoculum flask and spread onto the medium in the flask. Once growth has taken place in the inoculum flask (typically after a few days at 30 °C), the following operations are performed:

- Strain identity and traceability: ampoule number is registered
- Microbial purity: a sample from the inoculum flask is controlled microscopically for absence of microbial contaminants.

When sufficient amount of biomass is obtained and when the microbiological analyses are approved, the inoculum flask can be used for inoculating the seed fermenter.

A.4.1.4. The seed fermentation

The raw materials for the fermentation medium are mixed with water in a mixing tank. The medium is transferred to the seed fermenter and heat-sterilised (e.g. 120 °C / 90 min).

The seed fermentation tank is inoculated by transferring aseptically a suspension of cells from the inoculum flask.

The seed fermentation is run aerobically (sterile airflow), under agitation. The overpressure is kept above 0.2 atmosphere at all times, to prevent contamination.

Once a sufficient amount of biomass has developed, microbiological analyses are performed to ensure absence of contamination. The seed fermentation can then be transferred to the main fermentation tank.

A.4.1.5. The main fermentation

The raw materials for the medium are mixed with water in a mixing tank. The medium is transferred to the main fermenter and heat-sterilised (e.g. 120 °C / 60 min). If necessary, the pH is adjusted after sterilisation, with sterile pH adjustment solutions.

The fermentation in the main tank is run as normal submerged fed-batch fermentation.

The main fermentation is run aerobically (sterile airflow), under vigorous agitation. The overpressure is kept above 0.2 atmosphere at all times, to prevent contamination. The fermentation is run at a well-defined temperature.

Fresh medium is added aseptically when the pH increases above its set point, and the dissolved oxygen concentration rises. The feed rate is adjusted so that there is no accumulation of carbohydrates.

Other parameters are measured at regular intervals

- refractive index
- enzyme productivity
- residual glucose
- residual ammonia

Samples are also taken at regular intervals to check absence of microbial contamination.

A.4.2. Recovery

The recovery process is a multi-step operation designed to separate the enzyme from the microbial biomass and partially purify, concentrate, and stabilise the food enzyme.

The steps of this process involve a series of typical unit operations:

- pre-treatment
- primary separation
- filtration
- concentration
- stabilisation and preservation

A.4.2.1. Raw materials for recovery

The raw materials typically used in the recovery process are as follows:

- Potable water
- Filter aids or pre-coats (e.g. diatomaceous earth or perlite)
- Acids and bases for pH adjustment (e.g. phosphoric acid, sodium hydroxide, potassium hydroxide, ammonia)
- Antifoam, if necessary (e.g. polypropylene glycol)
- Flocculants (e.g. anionic polyacrylamide, poly(aluminium hydroxy)chloride)
- Stabilisation (e.g. glucose, sucrose)
- Preservatives (e.g. potassium sorbate, sodium benzoate)

A.4.2.2. Pre-treatment

To facilitate the separation, flocculants are used in a pH-controlled process.

A.4.2.3. Primary separation

The cell mass and other solids are separated from the broth by well-established techniques such as pre-coat vacuum drum filtration or centrifugation. The precoat used in the filter and the filter aid used in the process is diatomaceous earth (diatomite or perlite).

The primary separation is performed at well-defined pH and temperature range.

A.4.2.4. Filtration

For removal of residual cells of the production strain and as a general precaution against microbial degradation, filtration on dedicated germ filtration media is applied. Pre-filtration is included when needed.

The filtrations are performed at well-defined pH and temperature intervals, and result in an enzyme concentrate solution free of the production strain and insoluble substrate components from the fermentation.

A.4.2.5. Concentration

Ultrafiltration and/or evaporation are applied for concentration and further purification. The ultrafiltration is applied to fractionate high molecular weight components (enzymes) from low molecular weight components and is used to increase the activity/dry matter ratio. Evaporation is used to increase the activity while maintaining the activity/dry matter ratio.

The pH and temperature are controlled during the concentration step, which is performed until the desired activity and activity/dry matter ratio has been obtained.

A.4.2.7. Stabilisation and preservation

The enzyme concentrate is stabilised and preservatives are added in order to prevent microbial degradation. The enzyme concentrate is stored at 0-30 °C.

A.4.2.8. Process control

Apart from the process controls performed during the various fermentation steps and described above, the following microbial controls are also performed.

Samples are withdrawn from both the seed fermenter and the main fermenter:
a) before inoculation
b) at regular interval during cultivation
c) before transfer/harvest

The samples during all steps are examined by:
a) microscopy

b) plating culture broth on a nutrient agar and incubating for 24-48 hours.

Growth characteristics are observed macroscopically and microscopically.

During the microbiological control steps, the number of foreign microorganisms should be insignificant. The fermentation parameters, i.e. enzyme activity, temperature and oxygen as well as pH are also monitored closely. A deviation from the normal course of the fermentation may signal a contamination.

If a significant contamination develops, the fermentation is terminated. The fermentation is regarded as “significantly contaminated” if two independent samples show presence of contaminating organisms after growth on nutrient agar.

Any contaminated fermentation is rejected for enzyme preparations to be used in a food grade application.

A.5. Specification for identity and purity

The glucoamylase enzyme product complies with the purity criteria recommended for Enzyme Preparations in Food, Food Chemical Codex, 11 ^{h} edition, 2018.

In addition to this, the glucoamylase enzyme product also conforms to the General Specifications for Enzyme Preparations Used in Food Processing as proposed by the Joint FAO/WHO Expert Committee on Food Additives in Compendium of Food Additive Specifications, available online at: http://www.fao.org/food/food-safety-quality/scientific-advice/jecfa/jecfa-additives/en/.

Analytical data for an unstandardised, representative batch of the glucoamylase enzyme preparation is shown in the table below. These data show compliance with the purity criteria of the specification.

Control parameter	Unit	Specification	Batch NAFR 361
Glucoamylase activity	AGU/g		524
Heavy Metals[a]	ppm	Max 30	3.2
Pb	ppm	Max 5	ND[b] (DL < 0.5)[c]
As	ppm	Max 3	0.103 (DL < 0.1)
Cd	ppm	Max 0.5	ND (DL < 0.05)
Hg	ppm	Max 0.5	ND (DL < 0.03)
Total viable count	/g	Not more than 50000	400
Total coliforms	/g	Not more than 30	10
Enteropathogenic *E. coli*	/25g	Not detected	ND
Salmonella	/25g	Not detected	ND
Antimicrobial activity		Not detected	ND
Ochratoxin A	ppm		ND (DL < 0.0003)
Fumonisin B2	ppm		ND (DL < 0.0005)
Production strain	/g	Not detected	ND

a) Heavy Metals = Σ of Ag, As, Bi, Cd, Cu, Hg, Mo, Ni, Pb, Sb, Sn
b) ND = Not Detected
c) DL = Detection Limit

The methods of analysis used to determine compliance with the specifications are enclosed (Appendix 3).

The glucoamylase enzyme preparation is available under the commercial name Attenuzyme Core as a single enzyme formulation. The preparation is standardised in glucoamylase units (AGU/g; Appendix 3.1). The preparation does not contain known food allergens. Further details are given in the Product Data Sheet (Appendix 2.2)

A.6. Analytical method for detection

The glucoamylase enzyme preparation is to be used in the food industry as a processing aid. This information is not required in the case of an enzymatic processing aid.

B. Information related to the safety of a chemical processing aid

Not applicable – this application does not concern a chemical processing aid.

C. Information related to the safety of an enzyme processing aid

C.1. General information on the use of the enzyme as a food processing aid in other countries

The enzyme is used as processing aid during processing of starch-containing raw materials in a range of countries, where there are no restrictions of the use of enzyme processing aids or where the enzyme is covered by country positive list or specific approval.

The safety of the glucoamylase preparation has been evaluated and confirmed by external expert groups, as follows:

- Denmark: The enzyme preparation was safety assessed resulting in the authorisation of the enzyme product by the Danish Veterinary and Food Administration.
- Brazil: The enzyme was evaluated, approved and included in the Brazilian positive list – RDC 26/2009.
- Canada: The enzyme was evaluated by the Canadian authorities, Health Canada, and added to the Canada Food and Drug regulations, Division 16 Food Additives, Section B. 16.100, Table V, Food Additives that may be used as Food enzymes.
- China: The enzyme is included in the Chinese positive list – GB2760-2015.
- France: The enzyme is included in the French positive list for processing aids, including food enzymes (The French order of October 19, 2006 on use of processing aids in the manufacture of certain foodstuff), as amended.
- Mexico: Based on a dossier submitted by Novozymes, the Mexican food authorities, COFEPRIS, have approved the enzyme.

C.2. Information on the potential toxicity of the enzyme processing aid

(a) Information on the enzyme's prior history of human consumption and/or its similarity to proteins with a history of safe human consumption

A wide variety of enzymes are used in food processing. Enzymes including glucoamylase have a long history of use in food (Pariza and Foster, 1983[2] and Pariza and Johnson, 2001[3]).

Since the 1960s glucoamylases have been used extensively in various industrial food applications for the hydrolysis of starch to fermentable sugars, with major application in the starch, distilling, brewing, and baking industry (Godfrey, 1983[4]; Janda, 1983[5]; Poulson, 1983[6]; Reichelt, 1983[7]; van Oort, 2010[8]). Glucoamylase enzyme preparations from various sources are widely authorised in, e.g. Australia and New Zealand, Brazil, Canada, China, Denmark, France, Japan, Mexico.

(b) Information on any significant similarity between the amino acid sequence of the enzyme and that of known protein toxins

A sequence homology assessment of the glucoamylase enzyme to known toxins and allergens was conducted. No significant homologies to toxins were found. The complete search report is enclosed in Appendix 5.1.

Furthermore, safety studies as described below were performed. A summary of the safety studies is enclosed in Appendix 5.3.

The assessment of systemic toxicity (13-week oral toxicity study in rats) was performed on an enzyme concentrate of the same glucoamylase, produced by a previous production strain in the same strain lineage designated the BO-1 lineage.

This deviation is justified by the fact that the production strain belongs to Novozymes safe strain lineage of *Aspergillus niger*, as described in the following.

Aspergillus niger is recognised as a safe production organism with a long history of safe use for food ingredients (Section D).

As shown in Appendix 6.1 (Safety for humans and animals), the *Aspergillus niger* production strain has been developed in the BO-1 strain lineage which has been used for production at Novozymes A/S for more than 30 years. The BO-1 strain lineage has given rise to a number of food enzyme production strains, which are used for production of previously evaluated and regulatory approved food enzymes.

For six production strains from the BO-1 strain lineage, including the previous production strain used to produce the same food enzyme (represented by the tox test batch PPY 24900) as the glucoamylase object of this application, we have carried out comprehensive safety studies comprising at least the following: *in vitro* test for gene mutations in bacteria (Ames); *in vitro* test for chromosomal aberration or *in vitro* micronucleus assay; 13 week subchronic oral toxicity study in rats. The conclusions of these studies were in all cases favourable as described in Appendix 6.1.

These safety studies comprise different food enzymes manufactured from the same predecessor strain, supporting evidence that the BO-1 strain lineage is safe independently of the product manufactured. They also comprise studies on the same enzyme in different predecessor strains, supporting evidence that the controlled GM steps performed in the strain lineage do not cause safety issues.

Based on the above, the BO-1 strain lineage is considered to be a safe strain lineage, and therefore the use of food enzymes produced by strains from this lineage is considered safe.

Safety of the inserted DNA
As shown in Section E.1.2 (Introduced DNA) as well as Section 6.1.5 (Description of the production organism) in Appendix 6.1, the DNA insert does not contain sequences of concern.

Safety of the active enzyme
Glucoamylases, including from *Aspergillus niger*, have been used and approved for decades in many countries, among others in Denmark.

Furthermore, possible interactions of the glucoamylase activity with food constituents are considered of no safety concern.

The toxicological testing of the glucoamylase was conducted on two glucoamylase concentrates, i.e. tox test batches PPY 32789 and PPY 24900 which were produced according to the description given in Section 3.3.2 A.4. The tox test batches were glucoamylase enzyme concentrates without any addition of additives or other standardisation or stabilisation ingredients.

The tox test batch PPY 32789 was produced by the production strain currently used in production (described in this application). The tox test batch PPY 24900 was produced by a previous production strain in the same strain lineage (see C.2 b and Appendix 6.1).

The following studies were performed:

- Ames test for mutagenic activity (Appendix 5.4)
- *In vitro* micronucleus test (Appendix 5.5)
- Subchronic (13 week) oral toxicity study in rats (Appendix 5.6)

The main conclusions of the safety studies can be summarised as follows:

- Ames test: The results of the experiments give no indication of mutagenic activity of the glucoamylase tox test batch, PPY 32789, in the presence or absence of metabolic activation, when tested under the conditions employed in this study.
- *In vitro* micronucleus test: The glucoamylase tox test batch, PPY 32789, did not induce micronuclei in cultured human peripheral blood lymphocytes, either in the absence or presence of S-9 mix, under the experimental conditions employed for this study. Concentrations were tested up to 5000 microgram/mL, the recommended regulatory maximum for *in vitro* cytogenetic assays.
- Subchronic oral toxicity study: Oral administration of the glucoamylase tox test batch, PPY 24990, to rats at dosages of up to 10.0 mL/kg body weight/day for 13 weeks resulted in no treatment-related adverse effects. Consequently, the NOAEL was considered to be 10 mL/kg/day (equivalent to 5528 AGU/kg body weight/day or 1470 mg TOS/kg body weight/day).

Based on the present toxicity data it can be concluded that the glucoamylase enzyme preparation, represented by the two batches PPY 32789 and PPY 24900, exhibits no toxicological effects under the experimental conditions described.

C.3. Information on the potential allergenicity of the enzyme processing aid

(a) Information of the source of the enzyme processing aid

The glucoamylase enzyme is produced by an *Aspergillus niger* microorganism expressing the glucoamylase from *Talaromyces emersonii*. *Aspergillus niger* is ubiquitous in the environment and in general considered as a non-pathogenic fungus (see Section 3.3.2 D).

(b) Analysis of similarity between the amino acid sequence of the enzyme and that of known allergens

Enzymes have a long history of safe use in food, with no indication of adverse effects or reactions. Moreover a wide variety of enzyme classes (and structures) are naturally present in food.

The allergenicity potential of enzymes was studied by Bindslev-Jensen et al (2006[9]) and reported in the publication: "Investigation on possible allergenicity of 19 different commercial enzymes used in the food industry". The investigation comprised enzymes produced by wild-type and genetically modified strains as well as wild-type enzymes and protein engineered variants and comprised 400 patients with a diagnosed allergy to inhalation allergens, food allergens, bee or wasp. It was concluded from this study that ingestion of food enzymes in general is not likely to be a concern with regard to food allergy.

Additionally, food enzymes are used in small amounts during food processing resulting in very small amounts of the enzyme protein in the final food. A high concentration generally equals a higher risk of sensitisation, whereas a low level in the final food equals a lower risk (Goodman et al, 2008[10]).

A sequence homology assessment of the glucoamylase enzyme to known toxins and allergens was conducted (Appendix 5.1). The glucoamylase was compared to allergens from the FARRP allergen protein database (http://www.allergenonline.org)[11] as well as the World Health Organisation and International Union or Immunological Societies (SHO/IUIS) Allergen

Nomenclature Sub-committee (http://www.allergen.org)[12] using the recommended allergen method by EFSA.

Allergen risk assessment analysis of the glucoamylase was performed according to the EFSA scientific opinion using *allergen online*[11] and *allergen.org*[12] databases. The analyses of the glucoamylase's sequence identified one fungal allergen having an identity with a known allergen above the threshold of 35 % across an 80 amino acid window (Appendix 5.1).

The glucoamylase Sch c 1 of *Schizophyllum commune* was revealed to have up to 65.8 % identity with the glucoamylase produced by *Aspergillus niger* across an 80 amino acids window using the *allergen.org*[12] database.

Similarities with other allergens were below 35 %. Since it is generally accepted that proteins with such a low identity rarely share epitopes they were not considered in this report (Pearson, 2000[13]; Aalberse et al., 2001[14]; Hileman et al., 2002[15]; Ladics et al., 2007[16]).

Ladics and colleagues[16] compared the 80 amino acid sliding window search method to the conventional homology sequence comparison, the latter being the overall sequence homology between two proteins over the full stretch of the protein. Results showed that the conventional FASTA alignment was a more reliable search approach when testing sequence homology to known allergens using 35 % as the threshold in both methods. Conventional FASTA yields a much lower number of false positive, but the same number of false negatives (Ladics et al., 2007[16]). Because of the high number of false positives in the 80 amino acids sliding window search approach with a sequence homology threshold of 35 %, more and more experts suggest that this threshold may be too conservative and will overpredict cross-reactivity, which may result in the blocking of potentially useful products. There seem to be consensus that the risk of cross reactivity is considered to be very low beneath 50 % homology, low to moderate between 50 % and 70 %, and high above 70 % homology (Aalberse et al., 2000[17]; Hileman et al., 2002[15]; Stadler et al., 2003 [18] ; Goodman et al., 2008[10]). In 2008, Goodman and colleagues**Error! Bookmark not defined.** suggested to raise the threshold of the 80 amino acid sliding window search to 50 % sequence homology.

On the basis of the available evidence it is concluded that oral intake of the glucoamylase is not anticipated to pose any food allergenic concern. Further details regarding the risk assessment can be found in Appendix 5.2.

C.4. Safety assessment reports prepared by international agencies or other national government agencies, if available

The certificates of approval of the glucoamylase enzyme preparation by the Danish authorities following their safety evaluation in accordance with guidelines from the European Food Safety Authority (EFSA) are enclosed as Appendix 2.3 (for baking applications) and Appendix 2.4 (for syrup, brewing, and distilling applications).

D. Additional information related to the safety of an enzyme processing aid derived from a microorganism

D.1. Information on the source microorganism

The glucoamylase enzyme is produced by an *Aspergillus niger* microorganism expressing the glucoamylase from *Talaromyces emersonii*. The host strain is a modified (protease deficient) *Aspergillus niger* strain (BO-1) derived from a natural isolate of *Aspergillus niger* C40-1. The BO-1 strain lineage has been used by Novozymes for more than 30 years and has given rise to a number of food enzyme production strains, which are used for production of previously evaluated and regulatory approved food enzymes.

The glucoamylase production strain is a non-pathogenic, non-toxigenic, genetically-modified *Aspergillus niger* strain. The production strain is marker-free, and it does not produce secondary metabolites of toxicological concern to humans as explained in Section E 1.3, Section A.5 and Appendix 6.1.

Detailed information on the construction of the glucoamylase production strain is provided in Appendix 6.1.

D.2. Information on the pathogenicity and toxicity of the source microorganism

Aspergillus niger is ubiquitous in the environment and in general considered as a non-pathogenic fungus.

Aspergillus niger is classified as a group 1 microorganism according to EU Directive 2000/54/EC of the European Parliament and of the Council of 18 September 2000 on the protection of workers from risks related to exposure to biological agents at work. A group 1 microorganism means one that is unlikely to cause human disease.

Aspergillus niger as a species has been used safely for the production of food ingredients (e.g. citric acid) and food enzymes world-wide for decades.

Schuster et al. (2002)[19] reviewed the safety of *Aspergillus niger* and describe it as having a very long history of safe industrial use, being widely distributed in nature, and being commonly used for production of food enzymes and citric acid.

Aspergillus niger has been used in the industry since 1919, for instance for the production of citric acid, which could be an ingredient of foods such as soft drinks, fruit juices and jams. The US Food and Drug Administration (FDA) has listed *Aspergillus niger* as a source of citric acid (21 CFR §173.280[b]).

The JECFA (Joint FAO/WHO Expert Committee on Food Additives) has evaluated enzyme preparations derived from *Aspergillus niger*. This body of experts determined that enzymes from this source do not constitute a toxicological hazard (WHO, 1990)[20].

Carbohydrase, pectinase, protease, glucose oxidase, catalase, lipase and lactase enzyme preparations from *Aspergillus niger* are included in the GRAS petition 3G0016 (filed April 12th, 1973) that FDA on request from the Enzyme Technical Association (ETA) converted into separate GRAS Notices11 (GRN 89, 111, 132). Based on the information provided by ETA, as well as the information in GRP 3G0016 and other information available to FDA, the agency did

[b] Code of Federal Regulations: Sec. 173.280 Solvent extraction process for citric acid. Available at http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=173.280 (last visited 17 May 2018).

not question the conclusion that enzyme preparations from *Aspergillus niger* are GRAS under the intended conditions of use. Analogous conclusions were drawn in GRAS Notices GRN 158, 183, 214, 296, 345, 402, 428, 510, 651, 657, 699, and 703 which all describe food enzymes produced by *Aspergillus niger* strains.

Overall, it can be concluded that *Aspergillus niger* is widely accepted as a non-pathogenic organism and has a long history of safe use in food and food enzyme production.

D.3. Information on the genetic stability of the source organism

The inserted recombinant DNA is genetically stable during fermentation, as the inserted DNA is integrated into the chromosome.

The genetic stability of the production strain was tested at large-scale fermentation. The strain stability during fermentation was analysed by Southern blotting. No instability of the strain was observed.

For a more detailed description of the strain construction and characteristics, please see Section 3.3.2 E.

E. Additional information related to the safety of an enzyme processing aid derived from a genetically-modified microorganism

E.1. Information on the methods used in the genetic modification of the source organism

This section contains summarized information on the modifications of the host strain, on the content and nature of the introduced DNA and on the construction of the final production strain, as well as the stability of the inserted gene. The detailed information is provided in the confidential Appendix 6.

E.1.1. Host organism

The host strain is a modified (protease deficient) *Aspergillus niger* strain (BO-1) derived from a natural isolate of *Aspergillus niger* C40-1. The BO-1 strain lineage has been used by Novozymes for more than 30 years and has given rise to a number of food enzyme production strains, which are used for production of previously evaluated and regulatory approved food enzymes. The taxonomic classification of the strain is as follows:

Name:	*Aspergillus niger*
Class:	Eurotiomycetes
Order:	Eurotiales
Genus:	*Aspergillus*
Species:	*niger*

The classification of *Aspergillus niger* BO-1 was confirmed by Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH.

The recipient strain used in the construction of the *Aspergillus niger* production strain was derived from the C40-1 parental strain through a combination of classical mutagenesis/selection and GM-steps. These steps included the inactivation of proteins involved in protein glycosylation and other strain improvements for product purity and stability.

E.1.2. Introduced DNA

The vectors pHUda81 and pHUda211, used to transform the *Aspergillus niger* recipient strain are based on the well-known *Escherichia coli* standard vector pUC19 (Vieira and Messing, 1987[21]). No elements of these vectors, including the ampicillin resistance gene (*amp*), are left in the production strain. Both vectors contain an expression cassette consisting of an *Aspergillus* promoter, the glucoamylase coding sequence and an *Aspergillus niger* terminator. The glucoamylase gene (*amgGT*) is a product from a PCR reaction from genomic DNA isolated from *Talaromyces emersonii*. Furthermore, the vectors contain expression cassettes to increase the copy number of the *Aspergillus niger* acid stable amylase *asaA* and to introduce two markers for the selection during the construction of the production strain (*amdS* and *pyrG)*.

E.1.3. Construction of the Recombinant Microorganism

The production strain was constructed from the recipient strain through the following steps:

1) Plasmid pHUda81 was integrated into random loci in the recipient strain. This lead to the integration of the *amgGT*, *asaA*, and *amdS* expression cassettes.
2) A transformant was selected and subsequently modified to disrupt the *oahA* gene, responsible for the production of oxalic acid.
3) Plasmid pHUda211 was randomly integrated into the resulting strain, leading to the integration of the *amgGT* and *pyrG* expression cassettes.
4) A transformant was selected and screened for high glucoamylase activity and used as the final production strain.

E.1.4. Antibiotic Resistance Gene

No functional antibiotic resistance genes were left in the strain as a result of the genetic modifications. The absence of these genes in the production strain was verified by Southern blot analysis using the relevant antibiotic resistance gene probes.

E.1.5. Stability of the Introduced Genetic Sequences

The presence of the introduced DNA sequences was also determined by Southern hybridisation to assess the stability and potential for transfer of genetic material as a component of the safety evaluation of the production microorganism. The transforming DNA is stably integrated into the *Aspergillus niger* chromosome and, as such, is poorly mobilised for genetic transfer to other organisms and is mitotically stable.

F. Information related to the dietary exposure to the processing aid

F.1. A list of foods or food groups likely to contain the processing aid or its metabolites

The glucoamylase preparation is used as a processing aid during the manufacture of starch-based products. Glucoamylases convert starch by removing D-glucose units in a stepwise manner from the non-reducing end of the substrate molecule to produce glucose for further processing to a wide range of products, such as syrup, distilled alcohol, beer, bread, and juices.

F.2. The levels of residues of the processing aid or its metabolites for each food or food group

The glucoamylase enzyme preparation is used at minimum levels necessary to achieve the desired effect and according to requirements for normal production following GMP.

The enzyme is used during five methods for processing raw materials containing starch.

- During starch processing to produce syrups, the glucoamylase degrades polysaccharides into glucose. The use of the enzyme leads to an increased glucose yield.
- In beverage alcohol (distilling) processes the glucoamylase is used in order to degrade gelatinised starch and dextrins into glucose and other fermentable sugars.
- During brewing and other cereal based beverage processes the glucoamylase converts starch into fermentable sugars for the production of beer.
- In baking and other cereal based processes, the glucoamylase hydrolyses starch, from milled and damaged granules, into glucose which will be fermented by yeast.
- During processing of fruits and vegetables the glucoamylase degrades starch present in the fruits and vegetables in order to increase yield and to facilitate processing.

Use level
The enzyme preparation is used at minimum levels necessary to achieve the desired effect and according to requirements for normal production following GMP.

The conditions of use of the glucoamylase preparation during food processing do not only depend on the type of application, but also on the food production process of each individual food manufacturer. In order to ensure optimal effectiveness of the enzyme at an acceptable economic cost the dosage, reaction time, process conditions and processing steps are adjusted.

The highest dosage given for solid food is 3500 AGU per kg starch dry matter. This corresponds to 2.19 g of the glucoamylase enzyme preparation per kg starch dry matter equivalent to 928 mg TOS per kg starch dry matter (cf. Section 3.3.2 A.2.2.).

The highest dosage given for liquids (excluding distilled beverage spirits *vide supra*) is 4100 AGU per kg starch dry matter. This corresponds to 2.56 g of the glucoamylase enzyme preparation per kg starch dry matter equivalent to 1088 mg TOS per kg starch dry matter (cf. Section 3.3.2 A.2.2).

Enzyme residues in the Final Food
The glucoamylase preparation is used in processing of raw materials containing starch for the hydrolysis of starch to glucose. The enzyme is denatured by heat during processing or removed by distillation.

F.2.1. Estimates of human consumption

Method used for the dietary exposure assessment
An exposure assessment according to the Budget Method (Hansen, 1966[22]; Douglass et al., 1997[23]; ILSI, 1997[24]) has been performed, as the processed starch is used as an ingredient in a variety of food products and beverages.

Budget Method
Overall, the human exposure to the glucoamylase will be negligible because the enzyme preparation is used as a processing aid and in low dosages.

The Budget Method assumptions represent a "maximum worst case" situation of human consumption, in which the food enzyme object of the present application would be used at its maximum recommended dosages in all processed food and all processed beverages and not only in those food and drink processes described in Section F.2.

It is also supposed that the totality of the food enzyme will end up in the final food. This assumption is exaggerated since the enzyme protein and the other substances resulting from the fermentation are diluted or removed in certain processing steps.

As an example distilled beverage spirits will neither contain any TOS (Total Organic Solids) originating from the food enzyme preparation nor from the fermentation mash due to the distillation step(s).

Therefore the safety margin calculation derived from this method is highly conservative.

Assumptions in the Budget Method

Solid food	The maximum energy intake over the course of a lifetime is 50 kcal/kg body weight/day. 50 kcal corresponds to 25 g foods. Therefore, adults ingest 25 g foods per kg body weight per day. Assuming that 50% of the food is processed food, the daily consumption will be 12.5 g processed foods per kg body weight. It is further assumed that, in average, all processed food contains 25% starch (or starch-derived) dry matter = 3.12 g starch derived dry matter per kg body weight per day.
Liquids	The maximum intake of liquids (other than milk) is 100 ml/kg body weight/day. Assuming that 25% of the non-milk beverages is processed, the daily consumption will be 25 ml processed beverages per kg body weight.

	It is further assumed that all processed beverages contain 12% starch hydrolysates = 3.0 g starch derived dry matter per kg body weight per day. It is assumed that the densities of the beverages are ~ 1.

TMDI (Total amount of dietary intake) calculation
Starch-derived solid foods
The highest dosage given for solid food is 3500 AGU per kg starch dry matter, corresponding to 928 mg TOS per kg starch dry matter (cf. Section 3.3.2 A.2.2).

Based on this, 3.12 gram starch-derived dry matter in solid food will maximally contain:

928 mg TOS per kg / 1000 g per kg x 3.12 g = 2.90 mg TOS

Starch-derived liquids
The highest dosage given for liquids is 4100 AGU per kg starch dry matter, corresponding to 1088 mg TOS per kg starch dry matter (cf. Section 3.3.2 A.2.2).

Based on this, 3.0 gram starch-derived dry matter in liquids will maximally contain:

1088 mg TOS per kg / 1000 g per kg x 3.0 g = 3.26 mg TOS

Total TMDI of starch-derived solid foods and liquids

2.90 mg TOS + 3.26 mg TOS = 6.16 mg TOS

F.2.2. Safety Margin Calculation

The safety margin is calculated as dose level with no adverse effect (NOAEL) divided by the estimated human consumption (TMDI). The NOAEL dose level in the 13 weeks oral toxicity study in rats was concluded to be 1470 mg TOS/kg bw/day (cf. Section 3.3.2 C 2).

The estimated human consumption is 6.16 mg TOS/kg/day

The safety margin can thus be calculated to be 1470/6.16 or approximately **239**.

F.3. For foods or food groups not currently listed in the most recent Australian or New Zealand National Nutrition Surveys (NNSs), information on the likely level of consumption

Not relevant.

F.4. The percentage of the food group in which the processing aid is likely to be found or the percentage of the market likely to use the processing aid

In the estimate on human consumption given in F.2.1 above, it is assumed that all raw materials containing starch are processed using the glucoamylase object of this dossier as a processing aid at the highest recommended dosage.

F.5. Information relating to the levels of residues in foods in other countries

As described in F.2.1 above, a “worst case” calculation is made assuming that all organic matter originating from the enzyme is retained in the processed food product. The dietary exposure is estimated using the Budget Method, as the processed starch is used as an ingredient in a variety of food products and beverages.

F.6. For foods where consumption has changed in recent years, information on likely current food consumption

No significant changes in recent years are observed.

LIST OF REFERENCES

[1] Aunstrup K (1979) *Production, Isolation, and Economics of Extracellular Enzymes in Applied Biochemistry and Bioengineering*, Volume 2, Enzyme Technology, Eds. Wingard, L.B., Katchalski-Katzir, E. and Goldstein, L, pp. 28-68

[2] Pariza, MW and Foster, EM (1983) *Determining the Safety of Enzymes Used in Food Processing*. Journal of Food Protection, 46:5:453-468

[3] Pariza, MW and Johnson, EA (2001) *Evaluating the Safety of Microbial Enzyme Preparations Used in Food Processing: Update for a New Century*. Reg. Tox and Pharm 33: 173-186

[4] Godfrey T (1983) *Brewing. Industrial Enzymology. The application of enzymes in industry* (Godfrey T, Reichelt J eds.) The Nature Press, New York, 221-259

[5] Janda W (1983) *Fruit Juice. Industrial Enzymology. The application of enzymes in industry* (Godfrey T, Reichelt J eds.) The Nature Press, New York, 315-320

[6] Poulson PB (1983) *Alcohol - Potable. Industrial Enzymology. The application of enzymes in industry* (Godfrey T, Reichelt J eds.) The Nature Press, New York, 170-178

[7] Reichelt JR (1983) *Starch. Industrial Enzymology. The application of enzymes in industry* (Godfrey T, Reichelt J eds.) The Nature Press, New York, 375-396

[8] van Oort M (2010) *Enzymes in bread making. Enzymes in Food Technology* (RJ Whitehurst, M van Oort M, eds) Wiley-Blackwell, UK. 103-143

[9] Bindslev-Jensen C, Skov PS, Roggen EL, Hvass P, Brinch DS (2006) *Investigation on possible allergenicity of 19 different commercial enzymes used in the food industry*. Food Chem. Toxicol. 44, 1909-1915

[10] Goodman RE, Vieths S, Sampson HA, Hill D, Ebisawa M, Taylor SL, van Ree R (2008) *Allergenicity assessment of genetically modified crops—what makes sense?* Nature Biotechnology, 26 (1), 73-81

[11] Goodman RE, Ebisawa M, Ferreira F, Sampson HA, van Ree R, Vieths S, Baumert JL, Bohle B, Lalithambika S, Wise J, Taylor SL (2016). *AllergenOnline: A peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity*. Mol. Nutr. Food Res. 60(5):1183-1198.

[12] Radauer C, Nandy A, Ferreira F, Goodman RE, Larsen JN, Lidholm J, Pomés A, Raulf-Heimsoth M, Rozynek P, Thomas WR, Breiteneder H (2014) *Update of the WHO/IUIS Allergen Nomenclature Database based on analysis of allergen sequences*. Allergy, 69(4):413-419

[13] Pearson WR (2000). *Flexible sequence similarity searching with the FASTA3 program package*. In Bioinformatics Methods and Protocols (pp. 185-219). Humana Press, Totowa, NJ.

[14] Aalberse RC, Akkerdaas J, Van Ree R. (2001). *Cross-reactivity of IgE antibodies to allergens*. Allergy, 56(6), 478-490

[15] Hileman RE, Silvanovich A, Goodman RE, Rice EA, Holleschak G, Astwood JD, Hefle SL (2002). *Bioinformatic methods for allergenicity assessment using a comprehensive allergen database*. International Archives of Allergy and Immunology, 128(4), 280-291

[16] Ladics GS, Bannon GA, Silvanovich A, Cressman RF (2007). *Comparison of conventional FASTA identity searches with the 80 amino acid sliding window FASTA search for the elucidation of potential identities to known allergens*. Molecular Nutrition & Food Research, 51(8), 985-998

[17] Aalberse RC (2000) *Structural biology of allergens.* J. Allergy Clin. Immunol. 106, 228-238

[18] Stadler MB, Stadler BM (2003) *Allergenicity prediction by protein sequence*. FASEB J, 17, 1141-1143

[19] Schuster E, Dunn-Coleman N, Frisvad JC, van Dijck PWM (2002) *On the safety of* Aspergillus niger - *a review*. Appl. Microbiol. Biotechnol. 59, 426-435

[20] WHO (1990) *Enzyme Preparations*. Technical Report Series, No. 789, Thirty-fifth Report of the Joint FAO/WHO Experts Committee on Food Additive, World Health Organisation, Geneva. 15. http://apps.who.int/iris/bitstream/10665/37651/1/WHO_TRS_789.pdf (Last visited 17 May 2018)

[21] Vieira J, Messing J (1987) *Production of Single-Stranded Plasmid DNA*. Methods in Enzymology, 153, 3-11.

[22] Hansen SC (1966) *Acceptable Daily Intake of Food Additives and Ceiling on Levels of Use*. Fd. Cosmet. Toxicol. 4, 427-432

[23] Douglass JS, Barraj LM, Tennant DR, Wesley RL, Chaisson CF (1997) *Evaluation of the Budget Method for screening food additive intakes*. Food Additives and Contaminants. 14 (8), 791-802

[24] ILSI (1997) *An evaluation of the budget method for screening food additive intake.* Summary report prepared under the responsibility of ILSI Europe Food Chemical Intake Task Force, 1-12

Production, Isolation, and Economics of Extracellular Enzymes

Novo Research Institute
Bagsvaerd, Denmark

APPLIED BIOCHEMISTRY AND BIOENGINEERING
Volume 2

ISBN 0-12-041102-4

I. INTRODUCTION

Microorganisms synthesize numerous enzymes which all have their function in growth, metabolism, and autolysis. Most enzymes operate inside the cell in a protected and highly structured environment, but some enzymes are secreted from the cell. These extracellular enzymes are unique among microbial products, designed as they are to work for the benefit of the microorganism at a distance. Their function is to make food materials available to the microorganism by hydrolysis of high-molecular weight-compounds; hence they are all hydrolases.

Since the extracellular enzymes work in the medium surrounding the microorganism, outside the protecting cell membrane they must possess good stability to variations in the chemical and physical properties of this environment. Furthermore, the large volume in which the enzymes react makes it necessary for the microorganism to produce them in large quantities. These attributes make the extracellular enzymes suitable for industial utilization.

Extracellular enzymes have been a popular research area for biotechnologists; thousands of publications describing enzyme fermentation, properties, and applications have been issued in the last few years. For the nonspecialist it is difficult to evaluate what is of industrial relevance and how the enzyme industry operates.

It is the purpose of this chapter to present factual information about those extracellular enzymes that are now of commercial importance. Industrial production methods, enzyme properties, and application methods will be described, and an estimate of the economic importance of the enzyme industry will be given.

The information has been collected from scientific journals, patents, and brochures issued by the enzyme manufacturers. The information is often confusing or contradictory, and details about production methods are usually not disclosed. A critical evaluation has been attempted, but in order not to make the text unduly long, detailed discussions of doubtful points have, as a rule, been avoided.

II. HISTORY

Cultivation of microorganisms which produce extracellular enzymes is a simple way of obtaining hydrolytic enzymes in large quantities. In the Orient this method has been used since prehistoric times in the preparation of a variety of fermented foods and beverages. The principal enzymes were amylolytic and proteolytic, and the organisms used were predominantly of the genera *Aspergillus* and *Mucor*. Semisolid fermentation of moist rice was used, and production was on a small

scale with refinements based on generations of experience. These methods are still used for the manufacture of enzymes, but production has been concentrated on fewer, large producers, and the methods used in some factories are based on the latest scientific developments and extensive automation.

Commercial microbial enzymes were introduced to the West around 1890 when the Japanese scientist Takamine settled in the United States and started an enzyme factory based on the Japanese technology. The principal product was Takadiastase, a mixture of amylolytic and proteolytic enzymes prepared by cultivation of *Aspergillus oryzae*. Takamine improved the traditional process significantly by the replacement of rice with wheat bran.

Bacterial enzymes were developed in France in 1913 by Boidin and Effront, who found that the hay bacillus, *Bacillus subtilis*, produces an extremely heat-stable α-amylase when grown in still cultures on a liquid medium made by extraction of malt or grain. The enzyme was primarily used in the textile industry for removal of the starch size that protects the warp in the manufacture of cotton fabric. The protease also formed in this process was of minor interest but found some application in the tanning industry.

Around 1930, it was discovered that pectinases could be used with advantage in the preparation of fruit products. Crude enzyme mixtures were prepared in the same way as Takadiastase by strains of the *Aspergillus niger* group, e.g., *Aspergillus wentii*.

In the following years, several other hydrolytic enzymes were developed and sold on a commercial scale, e.g., pentosanase, cellulase, and lipase, but the technology was not satisfactory. The enzyme products were crude, of low activity, and rather expensive, thus, new applications progressed slowly and resulted in many disappointments.

After World War II, the fermentation industry underwent a rapid development as efficient methods for submerged cultivation in the production of antibiotics were discovered. These methods were soon adopted in the production of enzymes but did not result in extensive economic expansions in the industry.

Around 1960, glucoamylase was introduced as a means of hydrolyzing starch for the manufacture of dextrose. The enzyme had been known for a long time, but the usual preparations of *A. niger* gave incomplete hydrolysis. The discovery that removal of transglycosylase from the preparation allowed an almost 100% hydrolysis of starch to dextrose paved the way to industrial application. Today, enzyme hydrolysis has completely substituted acid hydrolysis in the manufacture of dextrose.

At the same time it was discovered that the alkaline protease from

Bacillus licheniformis, Subtilisin Carlsberg, was well suited for incorporation into detergents. Because the enzyme detergents were well accepted by consumers enzymes were very quickly added to almost all detergents, and this application became the largest outlet for microbial enzymes. After a setback in 1970 caused by some factory workers' allergic reactions to enzyme dust, dust-free preparations were developed and detergent enzymes have again become an important application of hydrolytic enzymes.

Substitution of microbial enzymes for calf rennet is an old dream of the industrial microbiologist. Numerous microbial enzymes have been tested for this purpose, but only three have wide application, viz., the proteases from *Mucor pusillus*, *Mucor miehei*, and *Endothia parasitica*. These enzymes were introduced in the 1960s and are now extensively used. When applied properly, they will produce excellent cheese; furthermore, they have the advantage over calf rennet that they are less expensive and their supply is unlimited.

In the last decade, development has been concentrated on enzymes with improved properties. An interesting example is the extremely heat-stable amylase from *B. licheniformis*, the first enzyme used at a temperature above 100°C (115°C). Furthermore, highly alkali-stable proteases from alkalophilic *Bacillus* species have been introduced. These enzymes have optimum activities up to pH 12 and are used with advantage in detergents and dehairing processes.

At present, industry is awaiting the next breakthrough, but it is not easy to see where it will happen. Extensive work is being done in a number of areas, e.g., in waste hydrolysis, especially of cellulose. Although cellulose is easily hydrolyzed by extracellular enzymes, an economical process is not available. Moreover, the natural product lignocellulose requires a chemical or enzymatic pretreatment which is not yet economical.

In conclusion it can be said that, although a large number of extracellular enzymes are known and may become of technical interest, there is no immediate large, new application in sight. From the technical literature one receives the impression that future developments will center on intracellular enzymes or immobilized microbial cells.

III. THE ENZYME INDUSTRY

A. Extracellular Microbial Enzymes for Industrial Use and the Producing Companies

The major enzyme producers are listed in Table I, and the most common enzyme products are found in Table II. A number of enzyme

TABLE I
COMPANIES PRODUCING EXTRACELLULAR ENZYMES FOR INDUSTRIAL USE

Denmark	1. Grindstedvaerket A/S, Brabrand
	2. Novo Industri A/S, Bagsvaerd
France	3. Soc. Rapidase, Seclin (subsidiary of No. 8)
Germany	4. Miles Kali-Chemie GmbH, Nienburg a.d. Weser (subsidiary of No. 20).
	5. Röhm GmbH, Darmstadt
Great Britain	6. ABM, Stockport, Cheshire
	7. Glaxo, Greenford, Middlesex
Holland	8. Gist Brocades NV, Delft
Japan	9. Amano, Nagoya
	10. Daiwa Kasei, Osaka
	11. Meiji Seika, Tokyo
	12. Meito Sangyo, Nagoya
	13. Nagase, Osaka
	14. Sankyo, Tokyo
	15. Shin Nippon, Tokyo
	16. Yakult Biochemicals, Nishinomiya
Switzerland	17. Swiss Ferment AG (subsidiary of No. 2)
	18. Schubert AG (subsidiary of No. 2)
United States	19. GB Fermentation Industries, Kingstree, South Carolina (subsidiary of No. 8)
	20. Miles, Elkhart, Indiana
	21. Novo Biochemical Industries, Frank Linton, North Carolina
	22. Pfizer, New York, New York
	23. Rohm and Haas, Philadelphia, Pennsylvania

suppliers without basic production and many enzyme products are not listed in the tables.

Because of a lack of information, no East-European enzyme producers have been included. It is known, however, that there is substantial enzyme production in the Soviet Union and some production in the German Democratic Republic, Czechoslovakia, Rumania, Hungary, and Bulgaria. Several large companies in Europe and the United States produce amylolytic enzymes for their own use. They are also excluded from the table. Novo and Gist Brocades are by far the largest producers in terms of turnover. Together, they account for the major part of sales of industrial enzymes.

B. Market Situation

World consumption of industrial extracellular microbial enzymes is difficult to evaluate, but an estimate expressed as the amount of pure active enzyme protein produced per year has been made (Aunstrup, 1977). Figure 1 shows the estimated world production in 1976.

TABLE II
COMMERCIAL EXTRACELLULAR ENZYMES, TRADE NAMES, AND SOURCES[a]

α-Amylase	*Aspergillus oryzae* Fungamyl (2); MKC clarase (4); Veron (5); Amylozyme (6); Mylex (7); Fermex, Mylase, Mycolase (8); Biodiastase (9); Asperzyme, Sanzyme, Takadiastase (14); Clarase, Dextrinase (20); Mycozyme (22); Diastase, Rhozyme (23)
	Bacillus amyloliquefaciens (*B. subtilis*) BAN, Aquazym (2); Rapidase (3); Optiamyl, Optimash, Optisize (4); Bacterase, Nervanase (6); Bactamyl, Maxamyl (8); Biokleistase, Kleistase, Diasmen (10); Biotex, Spitase (13); HT amylase, Tenase (20)
	Bacillus licheniformis Termamyl (2); Thermoase (10)
β-Glucanase	*Aspergillus* sp. Glucanase GV (1); β-glucanase 2000 D (6)
	Bacillus amyloliquefaciens Cereflo (2)
Cellulase	*Aspergillus* sp. Cellulase C (5); Cellase (8); Cellulase AP (9); Celluzyme (13)
	Trichoderma reesei (*T. viride*) Cellulase (2); Meicelase (11); Cellulase onozuka (16)
	Other sources MKC Cellulase (4); Cellulase P (5)
Dextranase	*Penicillin* sp. Dextranase (2)
Glucoamylase	*Aspergillus niger* (var.) AMG, SAN (2); Optidex, Optisprit (4); Ambazyme (6); Agidex (7); Amigase, Maxydrase (8); Diazyme (29)
	Rhizopus sp. Gluzyme (9); Sumyzyme (15)
Hemicellulase	*Aspergillus niger* Galactomannanase (2); MKC Hemicellulase (4)
Lactase	*Aspergillus niger* Lactase LP (8)
Lipase	*Aspergillus* sp. Lipase AP (9); Lipase B (23)
	Candida cylindraccae Lipase MY (14)
	Mucor miehei Piccantase (8)
	Mucor sp. Lipase AMP (9)
	Rhizopus sp. SNS (3); Saiken (13)
Mutanase	*Trichoderma* sp. Mutanase (2)

TABLE II (*Continued*)

Pectinase	*Aspergillus* sp. Pectolase GV (1); MKC Pectinase, Opticlar (4); Meliovin, Pectinol D, Rohament R (5); Klerzyme (8); Pectinex (17); Ultrazym (18); Spark L (20); Pectinil (23)
	Rhizopus sp. Macerozyme (16)
Protease	*Aspergillus niger* Proctase (11); Pamprosin (16)
	Aspergillus oryzae Veron P (5); Panazyme (6); Prozyme, Biozyme A (9); Sanzyme (14); Sumyzyme AP (15); Fungal Protease (20); Rhozyme (23)
	Rhizopus sp. Newlase (9)
	Bacillus amyloliquefaciens (*B. subtilis*) Neutrase (2); Rapidermase (3); Proteinase 18 (6); Protin (10); Bioprase, Nagase (16); Rhozyme (23)
	Bacillus licheniformis Alcalase (2); Optimase (4); Maxatase P (8)
	Bacillus thermoproteolyticus Thermoase, Thermolysin (10)
	Bacillus sp. *alkalophilic* Esperase, Savinase (2); Highly alkaline protease (8)
Protease, microbial rennet	*Endothia parasitica* Surecurd, Suparen (22)
	Mucor miehei Rennilase (2); Fromase (8); Marzyme (20); Morcurd (22)
	Mucor pusillus Emporase, Meito rennet, Noury lab (12)
Pullulanase	*Klebsiella aerogenes* Pulluzyme (6)

[a] Companies are identified by numbers referring to Table I.

World-wide sales of microbial enzymes amount to about 150 million U.S. dollars per year at present (1976). Table III shows the distribution on various enzyme types. It should be noticed that enzymes for medical and analytical applications are not included. From distribution of sales on industries (Table IV) it appears that the starch and detergent industries are predominant; furthermore, the latter is the only nonfood application of importance.

C. Enzyme Costs

Enzymes are always sold on an activity basis, but enzyme prices are difficult to compare because each manufacturer has his own method of analysis. Furthermore, the properties of enzymes from different

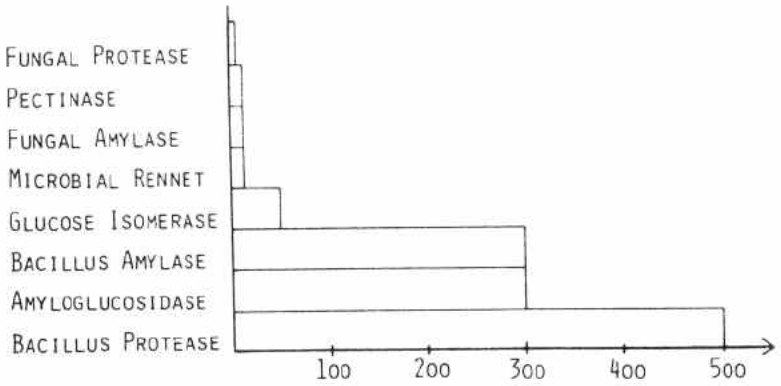


Fig. 1. Microbial enzyme production in tons of pure enzyme protein.

TABLE III
WORLDWIDE SALES OF ENZYMES[a]

Enzyme	Distribution (%)
Bacillus protease	35
Amyloglucosidase	14
Bacillus amylase	10
Glucose isomerase	14
Microbial rennet	5
Fungal amylase	4
Pectinase	10
Fungal protease	4
Others	4

[a] Distribution of enzyme types.

TABLE IV
DISTRIBUTION OF ENZYME SALES IN INDUSTRIES

Industry application	Distribution (%)
Starch	30
Detergent	35
Dairy	5
Distilling	5
Brewing	4
Fruit, wine	10
Milling, baking	5
Others	6

TABLE V
DIRECT ENZYME COST IN ENZYME PROCESSES: UNITED STATES, 1977

Application	Enzyme	Units	Enzyme cost (U.S. cents)
Washing	Protease	1 kg detergent	2–4
Starch			
liquefaction	Amylase	1 kg starch	0.2–0.5
saccharification	Glucoamylase	1 kg starch	0.4–0.8
isomerization	Glucose isomerase	1 kg starch	1.0–1.5
Cheese manufacture	Microbial rennet	1 liter milk	0.1
Alcohol manufacture	Amylase	1 liter alcohol	0.2–0.5
	Glucoamylase		0.7–1.4
Brewing	Amylase	1 liter beer	0.1
	Protease		
Baking	Amylase	1 kg flour	0.01
	Protease		
Juice	Pectinase	1 liter juice	0.1–0.2
Wine	Pectinase	1 liter wine	0.1–0.2
Dehairing	Protease	1 m^2 hide	1–5

sources often vary. This means that a comparison based on an analytical determination of the activity may be misleading because it is usually done under conditions which deviate from those of the actual application. The only proper way to compare enzymes from different sources is by their performance in the intended application.

The cost of an enzyme needed for a given process depends on the process conditions and, of course, on the factors that usually influence prices, such as customs, local trade patterns, and competition. Table V lists the actual enzyme costs of a number of enzyme applications. The enzyme costs are generally very small compared with the added value obtained by the enzyme process. Quality and reliability are therefore more important than a minimum price of an enzyme product.

IV. GENERAL METHODS FOR INDUSTRIAL PRODUCTION OF EXTRACELLULAR ENZYMES

Although enzymes are probably the most complex chemical compounds commercially available in large quantities, the methods used for the production of extracellular enzymes are in principle very simple: a microorganism is cultivated in a suitable medium from which the enzyme is subsequently recovered.

The problems lie in the details of the processes. To be competitive, the manufacturer must develop an economical and reliable production process which also meets today's strict requirements of safety and

hygiene. Appropriate modern equipment in good repair is a prerequisite for satisfactory results. The most important factor, however, is the optimum combination of a selected strain of microorganism and suitable fermentation conditions and recovery methods. This combination is achieved through comprehensive and time-consuming experiments. It is constantly improved and is the most important asset of the enzyme manufacturer.

As each company must develop its own methods without knowing the competitors' work, several approaches to the optimum result are likely. It would be interesting to compare the various methods used today, but this is unfortunately not possible. The information given in this chapter is limited to what has been published in scientific literature and patents.

A. Production Strain

In Table II it is shown that most commercial enzymes are produced by organisms belonging to the two genera *Bacillus* and *Aspergillus*. *Bacillus* species are well suited for enzyme production. Apart from the *B. cereus* group, including *B. anthracis*, they are harmless saprophytes which produce no toxins. They are easy to grow in high density and require no expensive growth factors. All *Bacillus* species form the heat-resistant endospores. Sporulation terminates the metabolism of the cell and is therefore undesirable in an enzyme production process. The spores are also undesirable because they may survive the recovery process and thus result in a high content of viable organisms in the final product. Fortunately, it is easy to prepare mutants that lack the spore-forming property. In many instances these mutants have a concomitant higher productivity of extracellular enzymes (Aunstrup and Outtrup, 1973). Many *Bacillus* species will produce polypeptide antibiotics, e.g., the production of bacitracin by *B. licheniformis*. These substances are usually not tolerated in food products; thus, it becomes necessary to use mutants that do not produce these substances. Extracellular enzyme formation by *Bacillus* has been extensively studied (Priest, 1977).

Aspergillus species may be regarded as the fungal parallel to the *Bacillus* species. The genus is highly variable and widespread, and very few species are regarded as pathogenic (*A. fumigatus*) or toxin-forming (*A. parasiticus*).

The *A. niger* and *A. oryzae* groups are the most frequently used for enzyme production (Table II). Based on morphological differentiation, the groups are divided into a number of species, but the differences between these are small. The *Aspergilli* are usually haploid in their

vegetative phase. This means that they are easily mutable; on the other hand, the mutants are rather unstable.

Since inoculation of fermenters or trays in surface cultivations is made with conidial spores, it is important to maintain the sporulating ability of the production strains. This may often be a problem because the strains with highly improved enzyme productivity generally have reduced ability to produce conidiospores.

Strain improvement is important in industrial laboratories. Since most of the organisms used are genetically unknown, it is difficult to use sophisticated genetic methods. When selecting high-yield mutants, there is often no better way than mutagenization and testing of each mutant in a shake flask culture. Each industrial microbiologist develops his own favorite tricks and advantageous methods, but there are no general rules.

It is the author's experience that microbiologists who have worked successfully with such problems for years often develop a "sixth sense" to detect valuable mutants. This is probably based on highly developed powers of observation combined with long-term experience. Highly mechanized testing systems are therefore not recommended as a substitute for traditional hand work for this type of development.

B. Fermentation

Two methods of fermentation are in general use: semisolid cultivation and submerged cultivation. Although the latter would seem to be more rational and more economical, a number of enzymes of fungal origin are still produced more economically by the former method.

1. Semisolid Cultivation

By this method the microorganism is grown on a particulate, moist substrate, preferably wheat bran with various additives. The method is preferred for the enzymes of Table VI.

The cultivation may be performed in trays with a substrate thickness of 1–10 cm or in rotating drums tumbling the substrate to ensure aeration. Combinations of growth forms may be used. It has been reported that, for instance, the use of a thick layer of substrate in part of the growth cycle (high heap) is an advantage.

The medium is prepared by mixing bran with water and additives and steam sterilizing the mixture in an autoclave equipped with a stirrer. The sterilized medium is transferred to the trays under aseptic conditions. Inoculation is made with spores in the autoclave after cooling, or in the cultivation equipment. Aeration is obtained by blowing

TABLE VI
ENZYMES PRODUCED IN SEMISOLID FERMENTATION

Enzyme	Organism
Glucoamylase	*Rhizopus* sp.
Pectinase	*A. niger*
Protease	*A. oryzae*
	A. niger
α-Amylase	*A. oryzae*
Lactase	*A. oryzae*
Rennet	*M. pusillus*

humidified air over the culture. To avoid overheating of fast-growing cultures, it may be necessary to use cooling surfaces. All handling of the substrate is mechanized as much as possible in a modern factory, and fully automated continuously operating factories are reported to exist. Unfortunately, no detailed descriptions of such factories are available.

Sterilization of the semisolid medium is difficult, and aseptic handling of the sterilized medium is almost impossible; thus, it is difficult to avoid minor contaminations. Fortunately, growth spreads slowly in the trays, and the massive inoculation decreases the ability of infections to overgrow the production organism. Nevertheless, infections are a major problem in this type of fermentation. Another problem is to prevent large numbers of spores of the production organism from spreading throughout the factory. Simultaneous production of different enzymes in the same factory is therefore difficult owing to cross contamination. Refinements like those used in the submerged fermentation technique, such as pH control or supplementation with medium components during the fermentation, are also complicated in semisolid fermentation.

The advantages of this system are the high aeration rate obtainable together with the low water concentration present. These advantages, and perhaps the fact that the mycelium is allowed to grow exposed to air, result in the formation of a broader variety of enzymes than would be possible in submerged culture.

2. *Submerged Cultivation*

The equipment and methods used in enzyme production are similar to the methods used in the preparation of antibiotics. Only the media and some fermentation conditions vary.

The fermentation medium should be inexpensive and support good growth of the microorganism. Usually, proteinaceous feed materials are used, combined with starch-containing materials such as grain and corn or carbohydrate-containing materials such as lactose, sucrose, or starch hydrolysates. Salts such as phosphates, magnesium salts, and ammonium salts or nitrate are added. Table VII gives some examples of media for enzyme production.

TABLE VII
TYPICAL FERMENTATION MEDIA

Submerged culture (composition in grams per liter)
Bacillus protease:
1. Starch hydrolysate 50, soybean meal 20, casein 20, Na_2HPO_4 3.3 (Churchill *et al.*, 1973).
2. Starch hydrolysate 150, lactose 4.3, cottonseed meal 30, brewers yeast 7.2, soy protein 3.65, K_2HPO_4 4.3, $MgSO_4$, H_2O 1.25, trace metals (Feldman, 1971).
3. Ground barley 100, soybean meal 30, pH adjusted to 9–10 with Na_2CO_3 (Aunstrup *et al.*, 1973).
Bacillus amylase:
4. Corn starch 40, ground corn (hominy) 100, corn steep liquor 65 (Smythe *et al.*, 1950).
5. Potato starch 100, ground barley 50, soybean meal 20, sodium caseinate 10, $Na_2HPO_4 \cdot 12H_2O$ 9 (Outtrup *et al.*, 1972).
Fungal protease:
6. Corn starch 30, corn steep liquor 5, soybean meal 10, casein 12, gelatin 5, distillers dried soluble 5, KH_2PO_4 2.4, $NaNo_3$ 1, NH_4Cl 1, $FeSO_4$ 0.01 (Lehmann *et al.*, 1977).
Fungal amylase:
7. Corn starch 24, corn steep liquor 36, NaH_2PO_4 47, $CaCl_2$ 1, KCl 0.2, $MgCl_2 \cdot 6H_2O$ 0.2 (Yamada and Tomoda, 1966).
Glucoamylase:
8. Corn starch 150, corn steep liquor 20, pH adjusted with gaseous NH_3 (Dworschak and Nelson, 1973).
Semisolid culture (composition in parts)
Lactase:
9. Wheat bran 100, 0.2 *N* HCl (containing traces of Zn, Fe, and Cu) 60 (Cayle, 1971).
Lipase:
10. Wheat bran 3, soybean meal 1, water 3 (Smythe and Drake, 1949).

The medium composition should be balanced so that the pH does not exceed the accepted limits during fermentation. This can be achieved by the addition of a buffer system such as phosphates or $CaCO_3$ to the medium. Another method is to add substances whose metabolism gives rise to a change in pH, such as salts of organic acids and nitrates that will raise the pH, or ammonium salts that will tend to lower the pH.

Glucose represses the formation of some enzymes, e.g., α-amylase. In these instances, it is important that the glucose concentration of the medium be close to zero at all times. This may be achieved by using slowly decomposable carbohydrates such as starch or lactose or by adding the glucose slowly during fermentation. The advantage of the latter method is that it allows the addition of large quantities of carbohydrates to the medium without producing a large concentration of dry substances which might tend to inhibit or retard growth.

Some enzymes, e.g., pectinase and lactase, require inducers for high-yield formation. These inducers—normally the enzyme substrate or their hydrolysates if the substrate is a polymer—must be added to the medium. If the inducer is expensive or unpleasant to work with, it is usually preferable to develop a mutant of the production strain that does not require the inducer.

The composition of the medium should not only be based on the requirement of the fermentation process, but the subsequent purification step must also be taken into consideration. It is important, therefore, that the composition of the medium be such that at the end of fermentation, total dry substance content and viscosity are low, the cell mass easily separable, and the content of free carbohydrate and amino acids in the liquid at a minimum.

Strict aseptic conditions are necessary in order to obtain a high enzyme yield and to ensure that no toxic substances or harmful microorganisms are introduced into the product. Aseptic conditions are particularly difficult to maintain in many enzyme fermentations performed at close to neutral pH and in rich media without any protecting antibiotic activity. This means that the equipment used must be of high standard and good steam sterilization of the medium is necessary.

All enzyme fermentations are aerobic, but in many instances they take place under oxygen-limiting conditions, e.g., amyloglucosidase (Aunstrup, 1977). In some of these cases the oxygen limitation appears to be an advantage, and an increased aeration rate will reduce the yield.

The composition of the broth at the end of fermentation depends on the initial medium composition. In percentage of initial dry matter

content of the medium, a typical composition is enzyme protein 1–5, residual nutrient and metabolites 5–10, cell mass 2–10.

C. Recovery and Finishing

The purpose of the recovery process is to prepare a finished product of satisfactory purity and stability in as few steps as possible and with a minimum loss of enzyme activity.

1. Pretreatment

An important point in the recovery process is to prevent introduction of toxic materials or harmful microorganisms into the product. When the broth leaves the fermenter, aseptic conditions can no longer be maintained; and, as the broth is a good substrate for a variety of microorganisms, precautions to avoid contamination must be taken. This is done by cooling the broth to about 5°C as quickly as possible, e.g., in a counter-current heat exchanger, and by using high-quality standard dairy equipment throughout the process. Furthermore, scrupulous cleanliness is maintained in the recovery plant. The use of chemical preservatives is normally not acceptable at this stage, first, because the high cell density requires a high dosage for the substance to become active, and second, because most of the effective preservatives are undesirable or prohibited in the final enzyme preparation if it is to be used as a food additive. The fermentation broth contains the enzymes in soluble form mixed with solids from the medium components and bacterial cells, fungal mycelium, or their decomposition products.

The solids are removed by filtration or centrifugation. They are often of colloidal nature and difficult to remove directly, in which case coagulating or flocculating agents may be added. An old device, which is often useful, is to precipitate calcium sulfate or calcium phosphate in the broth. Usually, it is more efficient to employ some of the flocculating agents which have been developed in recent years, e.g., for water treatment. They are polyelectrolytes such as polyamines, and the treatment may involve reaction with electrolytes of opposite charge, such as aluminum salts. In some cases it may also be necessary to add filter aid, e.g., in the form of diatomaceous earth before filtration.

The separation process is done on drum filters or leaf filters; if centrifugation is preferred, high-speed disk centrifuges are used. In many cases it is advantageous to mature the broth before separation, i.e., to let the broth set for some time so that the colloids aggregate and thus ease the separation process. However, economy and contamination risk limit this kind of operation.

2. Purification

There are various ways to prepare a commercial product from the pretreated broth. The most common methods are shown in Fig. 2.

The first step usually is to concentrate the enzyme by evaporation or ultrafiltration. Evaporation is performed in traditional or custom-built multistage vacuum evaporators. The liquid must be kept at a low temperature (30°–50°C). Nevertheless it is difficult to prevent activity losses of ~10% even for very stable enzymes. An additional disadvantage of this method is that the low temperature makes microbial growth possible, so that frequent cleaning of equipment and careful control of the product are necessary. Another problem is that the dry-substance content of the broth increases with evaporation, in some cases causing high viscosity and difficult handling.

In these instances ultrafiltration is an advantage. This technique has been developed to a high degree in recent years. Plant-scale equipment is readily available, the process is inexpensive, and it offers the additional advantage that substances of molecular weight below 10,000 are removed from the concentrate. Furthermore, the process can be performed at a low temperature (5°C), thus keeping the activity loss and contamination risk at a minimum. The disadvantages of this method are that the precipitate formed during the concentration process tends to clog the membranes. These membranes are often made of cellulose derivatives which are attacked by the cellulase enzymes present in many fermentation broths of fungal origin. In such instances other membrane types should be used, e.g., composite membranes based on polysulfones on a carrier fabric.

The concentrated enzyme solution is usually turbid due to precipitates formed in the concentration process. It also contains varying amounts of bacteria or other microorganisms. These are removed by a germ filtration process. Because of the small capacity of the germ filters, the process must be preceded by a polishing filtration. However, it should be noted that the previously very popular asbestos-based germ filters now have a limited application because of the health hazards of asbestos fibers. Instead, cellulose filters or membrane filters may be used.

Before the product is ready for sale, the clear filtrate must be mixed with stabilizers and preservatives, and the activity must be standardized to the specified value. Stabilizers are used to increase the storage stability of the enzyme preparation. Their composition varies from one enzyme to the next. Useful agents are salts, proteins, starch hydrolysates, and sugar alcohols.

The number of permitted preservatives is very limited because they should be both enzyme-tolerant and approved food-additive sub-

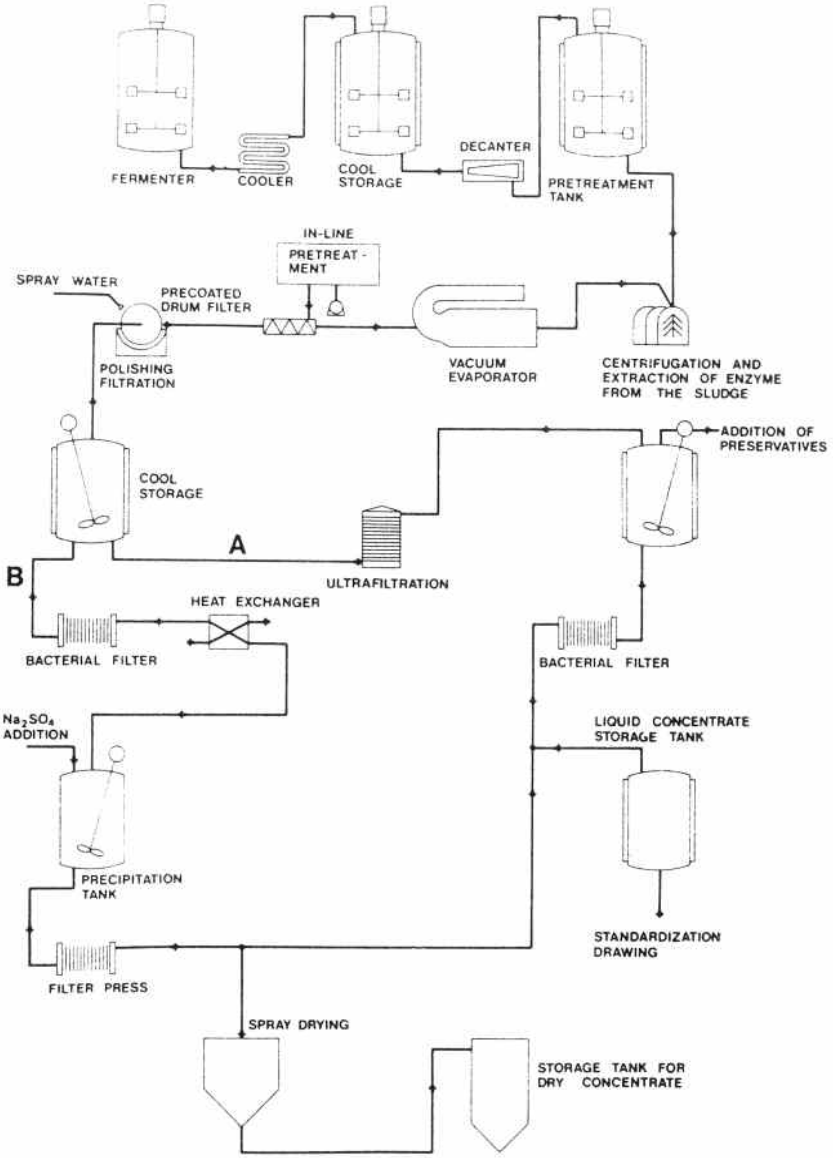


Fig. 2. Examples of enzyme recovery. (Courtesy of T. K. Nielsen, Novo Industri.)

stances. One of the best preservatives is NaCl in a concentration of 18–20%. Other, less efficient, agents are benzoate, parabene, or sorbate. Standardization and testing of the final preparation constitute an important part of the manufacturing process. Details follow later.

Liquid enzyme preparations like the one described are preferred because they are relatively inexpensive to make and safe and easy to use in most applications. In some cases, however, solid enzyme preparations must be used; e.g., for flour treatment or in granulated detergents.

To make a solid enzyme preparation, the simplest method is to spray-dry the clear concentrate. For good results, it is important that the concentration of low-molecular-weight substances, which would tend to make the product sticky or hygroscopic, be low. Consequently, preparations concentrated by ultrafiltration are better suited for spray-drying than those concentrated by evaporation. Spray-drying has some other disadvantages: the process is expensive (about 5–10¢/liter of water removed), the enzymes are subjected to fairly high temperatures and oxidizing conditions, impurities are not removed in the process, and finally the product has a low weight/volume ratio. An advantage is that spray-dried preparations usually are easily soluble.

Precipitation is preferred when some purification is desired in the process. Two methods are used by the industry: solvent precipitation and salt precipitation. Both methods have advantages and drawbacks. Salt precipitation is traditionally done with ammonium sulfate. Since this, however, is not acceptable in detergents, enzymes for this purpose are made with sodium sulfate. The method used is simple. Salt in dry form is added to the clear enzyme solution until the required concentration is obtained, and the precipitate is removed by filtration and dried.

For economic and environmental reasons the salts present in the supernatant are regenerated. Regeneration takes place in plants specially designed for recovery and crystallization of the salt. In this way it is possible to keep the costs of the salt precipitation process at a fairly low level. The advantage of the salt precipitation process is the rather small investment required and the good solubility of the product. The disadvantages are the high salt concentration present in the product, less efficient removal of impurities than in the solvent precipitation process, and the difficulty of maintaining aseptic conditions during the process.

Solvent precipitation may be performed with ethanol, acetone, or other water-soluble solvents. The solvent is added to the enzyme solution in the proper concentration and the precipitate is filtered off and dried. The supernatant solvent is purified and recovered by distillation.

This process is advantageous in that it results in a product of high purity and high activity. The product may be more difficult to dissolve than the salt-precipitated product. Investments in explosion-proof

equipment and recovery equipment are high, and the solvent loss must be kept at a very low level to operate economically.

The precipitation processes may be improved by various modifications such as fractional precipitation or multiple precipitation with intermediate purification steps. Precipitation with tannin, lignin, lignosulfonic acid, and other materials has been described in the literature, but these processes have had limited use.

In several instances, the purification process includes a step that removes an undesirable side effect. In principle, this can be done by any selective method of separation of inactivation. The problem is that removal of the undesirable component must be completed without substantial loss of the main component. An example of such a process is the removal of lipase in the production of microbial rennet by *M. miehei*. The lipase activity present in the broth after fermentation is reduced to less than 10% by leaving the broth at a pH below 3.5 for a few hours. The loss of rennet activity is insignificant (Schleich, 1971).

Another example is the removal of transglucosidase activity from glucoamylase preparations. This has been a serious problem in industry, and over the years many patent applications have been filed on methods to solve the problem. A number of methods are listed in Table VIII.

TABLE VIII
PATENTED METHODS FOR THE REMOVAL OF TRANSGLUCOSIDASE FROM GLUCOAMYLASE PRODUCTS

Year	U.S. Patent No.	Principle
1962	3,042,584	Adsorption on clay mineral
1962	3,047,471	Precipitation with lignin and tannic acid
1962	3,067,108	Precipitation with sulfonated compounds
1963	3,101,302	Adsorption to clay mineral with solvent
1963	3,108,928	Treatment at pH 9–10 and 22–55°C
1964	3,117,063	Precipitation with lignin or tannic acid and sulfonated compounds
1964	3,134,723	Precipitation with Ba, Ca, Sr, Cd, Pb, Mn, or Zn at pH 4.5–8.5
1965	3,268,417	Treatment with protease at pH 6–9
1967	3,303,107	Treatment at pH 1–3 and temperatures up to 70°C
1968	3,380,891	Treatment with sulfonated fatty acids at pH 1.5–4
1968	3,380,892	Treatment with cation-forming precipitate with oxalic acid at pH 1.5–3
1969	3,483,084	Treatment with chloroform
1969	3,483,085	Treatment with hetero poly(acid)s such as phosphomolybdic acid

3. *Finishing of Solid Enzyme Preparations*

The dried preparations from salt or solvent precipitation form large lumps which are unattractive to sell. The simplest way to prepare a marketable product is to grind the preparation to a fine powder in a mill and standardize the activity by the addition of inert substances such as salt, lactose, or other suitable substances. In the preparation of enzymes for flour treatment, standardization is usually done with flour.

Such a preparation will be satisfactory in many ways but it has one disadvantage: the particles are very small and the dust formation by handling will be relatively high. Since exposure to enzyme dust may cause allergic reactions, it is preferable to make the enzyme preparations in a way that minimizes the dust-forming ability. A number of such methods have been developed. The simplest one is admixture of dedusting compounds such as polyethylene glycol; another method is granulation with inorganic salt, e.g., by fluid bed agglomeration. The methods preferred at present involve granulation and coating with inert wax. In the so-called prilling process the enzyme is mixed with melted wax and spray-cooled. This forms granules of about 0.5–0.8 mm diameter in which the enzyme particles are uniformly distributed.

An improved process is shown in Fig. 3. The enzyme in this process is mixed with an inert filler (e.g., salt), a binder (e.g., carboxymethylcellulose), and water to make a paste. The paste is then extruded and shaped into spheres in a so-called marumerizer. After drying, the spheres are coated with a layer of wax material. A uniform particle size, e.g., 0.5–2 mm diameter, is obtained in this way, and dust formation is insignificant.

4. *Immobilization of Extracellular Enzymes*

This chapter will not describe the many methods developed to immobilize enzymes. Reference is made to Volume 1 of this series.

V. STANDARDIZATION AND CONTROL

The final step of the manufacturing process is standardization of the enzyme activity and control of other properties of the product. Manufacturers and customers have a common interest in the reliable performance of these tests, both for economic reasons and to ensure safe use of the product.

The number of tests has grown considerably in the last few years and will probably increase much more in the future. The manufac-

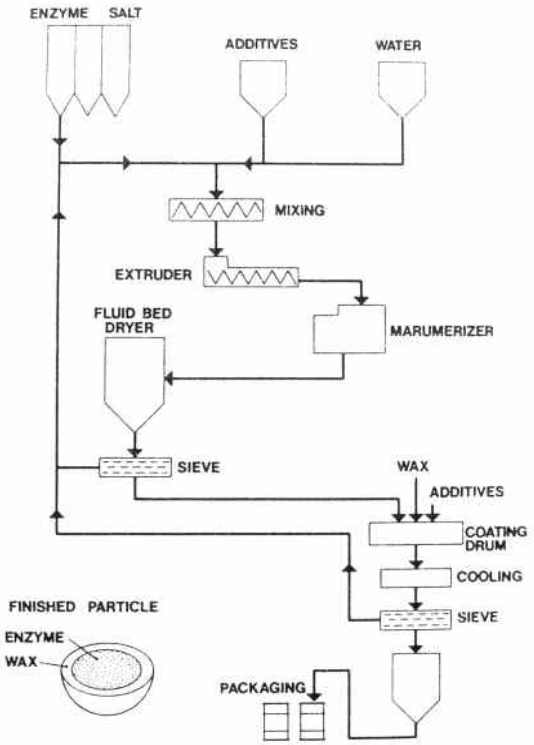


Fig. 3. Preparation of dust-free enzyme. (Courtesy of T. K. Nielsen, Novo Industri.)

turer, therefore, has to reckon with a substantial expense for this part of the manufacturing process, and he has to invest in extensive facilities for control laboratories.

A. Standardization

In principle, enzyme preparations are always sold on an activity basis. The manufacturer will promise or guarantee a certain effect in application, and the customer expects this effect to take place when he uses the product. He will also expect properties to be the same from batch to batch. To ensure that the enzyme concentration is maintained at the desired level, the manufacturer must determine the enzyme activity and standardize the product with inert material. Accurate and reliable analytical methods are necessary, and means are required to ensure that the activity level does not change in time.

The Commission on Biochemical Nomenclature has recommended defining enzyme activity as reaction rate, 1 catal being the amount of

enzyme which converts 1 mole of substrate per second. This definition can only be used if the substrate is well defined and the enzyme reaction known. Industrial enzymes, as a rule, will react with variable, natural substrates, and the enzyme reaction is often not known in detail. It has been necessary, therefore, to use other ways to define the enzyme activity. The following example illustrates this.

A glucoamylase unit may be defined as the amount of enzyme that hydrolyzes starch at such a rate that 1 mole of glucose per second is released under specified reaction conditions. The method will give reliable and reproducible results as long as the same batch of starch and enzyme is used. However, experience proves that the reaction rate depends on the starch quality, and the test is sensitive to the presence of α-amylase, a side activity of all glucoamylase preparations. Consequently, this method cannot be employed for standardization. A simple, well-defined substrate like maltose, which is not hydrolyzed by α-amylase, can solve the problem. This method also has drawbacks: pure maltose is difficult to obtain and maltose preparations are often contaminated by maltotriose which is hydrolyzed at a higher rate. Furthermore, glucoamylase is not used for maltose hydrolysis; hence the method cannot be applied in comparing glucoamylases of *Aspergillus* and *Rhizopus*, which have different ratios between hydrolysis of starch and maltose.

The problems described are simple compared with the analytical difficulties of other important enzymes, such as milk-coagulating enzymes, proteases, and pectinases. In all cases, the substrate is variable, side activities are important, and enzyme products of different origin do not react in the same way to changes in reaction conditions.

It is unlikely that standard methods applicable to all enzyme preparations and with an absolute unit definition will be developed. Instead, enzyme manufacturers must develop their own analytical methods based on reactions which they believe to be reliable and relevant to the application of their enzyme. To overcome the variation in substrate, it is customary to introduce enzyme standards in the analytical methods and to adjust the result accordingly. In this way it is possible in most cases to standardize the enzyme preparations with a variance of less than 10%.

B. Control

New enzyme products must be thoroughly tested to ensure their safety in the application intended. For use in food production new enzymes must be approved by relevant government bodies; in the United States it is the Food and Drug Administration (FDA). The

production process must comply with "good manufacturing practices," and a representative sample of the product is put through thorough toxicological tests. This is an expensive and time-consuming process (2–5 years). Requirements of other applications vary, but enzymes such as detergent enzymes that may come into contact with consumers must undergo a test program as rigid as that for food enzymes.

Routine control of enzyme products involves the following:

- tests appropriate to the application, such as the absence of transglucosidase in glucoamylase or the determination of the protease level in amylase;
- storage stability tests and analysis of the concentration of preservatives and stabilizers such as Ca^{2+};
- appearance tests of color, odor, amount of precipitate, and particle size;
- tests for undesirable impurities such as heavy metals, mycotoxins, and antibiotic activity;
- determination of enzyme dust level in granulated enzyme preparations;
- microbiological tests which usually involve a total viable count and tests for absence of pathogens or indicator organisms such as coliforms. A test for absence of production organism is generally included.

VI. PROTEOLYTIC ENZYMES

Extracellular proteases are probably the most widespread microbial enzymes. They are easy to detect and isolate, and they often occur in large concentrations. Consequently, they have been popular research objects, and there is an extensive literature on production and properties of proteases from various microorganisms.

Proteases with a wide variety of specificities are known, but very few are produced industrially. The most important are the alkaline serine proteases and neutral metalloproteases, which are formed by *Bacillus* species and the acid proteases from *Aspergillus* and *Mucor*.

A. Serine Proteases

Serine proteases owe their name to the fact that serine is an essential amino acid of the active center. They are endoproteases, and they have an alkaline pH optimum. The microbial serine proteases are usually very stable and have a strong proteolytic activity with a low specificity. Serine proteases are inhibited by diisopropyl fluorophosphate or phenylmethylsulfonyl fluoride.

1. Subtilisin Carlsberg

This enzyme was discovered in 1947 by Linderstrom Lang and Ottesen at the Carlsberg Laboratory. Subsequent work at this laboratory and by others has resulted in an extensive knowledge about the composition and properties of the enzyme. The amino acid sequence and tertiary structure of the enzyme are known.

The commercial utilization of the enzyme started in 1960 when it was discovered that the enzyme had excellent properties for use in detergents. Subtilisin Carlsberg is now the most widely used detergent protease.

Subtilisin Carlsberg is produced by *B. licheniformis*. Due to the uncertainty of *Bacillus* taxonomy at the time the enzyme was discovered, the producing organism was named *B. subtilis*. Unfortunately, this name is still erroneously used by many authors.

a. Production Method: A simple and good laboratory method for production of the enzyme was developed at an early stage (Güntelberg, 1954). The organism was grown on glucose, casein hydrolysate, and salts; and it was demonstrated that yields increased if glucose was fed during the fermentation. The organism has no growth requirements, but it is necessary to include organic nitrogen sources in the medium to obtain a high yield. Many strains of *B. licheniformis* will produce protease in high yields, for instance, the strain NRRL B 3723 (Feldman, 1971).

In commercial production the organism is grown at temperatures between 30° and 40°C and at neutral pH. A concentrated medium with a high content of protein or protein hydrolysate is used (Table VII), and it may be an advantage to feed carbohydrate during the fermentation.

The enzyme production starts when the maximum cell count is reached after 10–20 hr of growth and continues at an almost constant rate throughout the fermentation period if the fermentation is run properly. Laboratory experiments with production in continuous fermentation have been reported (Jensen, 1972). The continuous production process does not present special problems; whether it is feasible or not is entirely a question of economy.

When starch hydrolysates are used as a carbon source the α-amylase and α-glucosidase of *B. licheniformis* hydrolyze both α-1,4- and α-1,6-glucosidic bonds so that the starch is completely metabolized. However, the level of these enzymes is usually insignificant in the broth. The protease hydrolyzes all proteins of the medium, and at the end of the fermentation period the protease is practically the only protein dissolved in the broth. The yield is usually high. It is possible

to obtain a yield of enzyme protein of over 10% of the initial protein content of the medium.

The enzyme is primarily sold in the form of dust-free granulates for detergent use. Commercial granulates contain 1–5% active enzyme protein. The enzyme is also stable in liquid form, and the enzyme content of liquid preparations is about 2%.

b. Properties and Application: Subtilisin Carlsberg is an alkaline protease (pH optimum in hydrolysis of casein about 10). Over 80% of the activity is maintained in the pH range 8–11. The enzyme is stable at pH 5–10 at 25°C, and up to ~50°C for 1 hr at pH 8.5. The enzyme does not depend on Ca^{2+} for stability and will maintain stability in the presence of sequestering agents such as tripolyphosphates and ethylenediaminetetraacetic acid. Inactivation of the enzyme is rapid at pH values below 4 or above 11.5 and at temperatures above 70°C. Oxidizing agents such as hypochlorite and hydrogen peroxide destroy the enzyme rapidly, but stabilized hydrogen peroxide, such as in sodium perborate, does not harm the enzyme activity. The specificity is broad; only 30–35% of the peptide bonds in casein will not be hydrolyzed by the enzyme.

Subtilisin Carlsberg has performed excellently in detergents for many years. The optimal enzyme concentration in most detergents is ~0.015–0.025% active enzyme protein. The pH of the suds is usually ~9, which is close to the optimum of the enzyme. The proteolytic action prevails during the washing process up to 55°–60°C. In this temperature range heat inactivation sets in, and it coincides with the temperature at which H_2O_2 is released in perborate-containing detergents.

Proteases are obviously useful for washing of clothes stained with blood or other proteinaceous matter. It has also been proven that proteinases improve the general washing efficacy because proteinaceous impurities make dirt adhere to the fabric.

The use of detergent enzymes depends on washing habits. They are more widely used in Europe where perborate and high-temperature wash bring out the advantages of detergent enzymes. Because Subtilisin Carlsberg is a durable and inexpensive protease, it may be used in many fields where proteolysis under alkaline conditions is needed, for instance, in preparation of protein hydrolysates for food or feed purposes.

2. *Subtilisin Novo (Subtilisin BPN)*

Subtilisin Novo or Subtilisin BPN is produced by *Bacillus amyloliquefaciens*. This new species is separated from *B. subtilis* (Welker

and Campbell, 1967), because the strains used for industrial enzyme production have many properties in common, primarily the properties of their extracellular enzymes which were different from those of the neotype of *B. subtilis*, the Marburg strain.

B. amyloliquefaciens has been used for industrial enzyme production for over 50 years. Practically all protease preparations of bacterial origin produced before 1960 were made from this organism. At present, Subtilisin Novo is only used to a minor extent. It is the proteolytic side activity of the amylase preparations that are used for some detergents.

a. Production Methods: Most strains of *B. amyloliquefaciens* will produce protease in good yields. Welker and Campbell (1967) have listed a number of strains that are suitable for production of the enzyme. Medium and growth conditions are similar to those used for *B. licheniformis*, but the protease yield is usually somewhat lower.

During fermentation several other extracellular enzymes will be produced; primarily α-amylase, β-glucanase, neutral proteinase, and hemicellulase. The latter two enzymes are unstable and are usually present in low concentrations in the final preparation if no special precautions are taken.

Methods for removal of the α-amylase activity by fractional precipitation with calcium acetate and solvents (Keay and Anberson, 1971) or by oxidation with hypochlorite (Hoerle, 1976) have been described but have found only limited commercial application. Commercial preparations are usually made in liquid form, as dedusted powders or dust-free granulates for detergents. As a rule, they contain large amounts of α-amylase. The content of Subtilisin Novo is generally less than 1%.

b. Properties and Application: Subtilisin Novo is closely related to Subtilisin Carlsberg, and this is reflected in the stability and activity. In short, the temperature and pH range is a little narrower for Subtilisin Novo than for Subtilisin Carlsberg. The enzyme has a limited application—mainly in combination with α-amylase in detergents.

3. *Proteases from Alkalophilic Bacillus Species*

In 1967, it was discovered that *Bacillus* strains that grow at pH values over 10 produce proteolytic enzymes that are active and stable at pH values up to 12 (Aunstrup *et al.*, 1972). A large number of strains were isolated and several different proteases were found. All were serine proteases, and, apart from their better stability at high pH values, the properties were similar to the Subtilisin-type proteases. The molecular basis of the good alkali stability has not been subjected to

intensive studies, but it is characteristic that the isoelectric points of these enzymes are around pH 11 (as compared to pH 9.7 for Subtilisin Carlsberg).

a. Production Methods: Preparation methods for the proteases have been described in several patents (Aunstrup *et al.*, 1973; Horikoshi and Ikeda, 1977), in which suitable strains are also mentioned. The production methods are similar to those used for making Subtilisins, but the fermentation process is characteristic in that pH must be kept above 7.5 at all times. Otherwise the culture will die and lyse.

The high pH may be obtained by addition of alkali such as sodium carbonate, or the medium may be prepared with salts of metabolizable acids such as lactates. Some of the strains are amylolytic so that starch hydrolysates may be used as the carbon source. The strains may secrete small amounts of other alkalophilic enzymes, but the proteases will usually constitute over 90% of the enzyme content of the broth.

The enzymes are prepared in granulated form for detergents and as dedusted powder for dehairing. The protease content of the preparations is generally from 1 to 2% active enzyme protein.

b. Properties and Application: The new proteases are stable and active in the pH range 6–12 and may be used at temperatures up to 60°C. In detergents they are generally superior to Subtilisin Carlsberg. This is particularly pronounced if the alkalinity is high or the traditional sequestering agent, tripolyphosphate, is replaced by other sequestering agents such as citrate or gluconate.

The proteases are very useful in the dehairing process. It is customary to swell the hides with lime and loosen the hair with sodium sulfide. This process is effective, but highly undesirable for ecological reasons and because of the risk involved in handling sulfides. A satisfactory result may be achieved when these alkaline proteases are used in combination with lime. However, the application has so far been limited mainly for economic reasons.

B. Metalloproteases

The metalloproteases or neutral proteases have received much less attention than the Subtilisins, scientifically as well as commercially, although they too are widely distributed in nature. The metalloproteases are endoproteases: Zn^{2+} is essential to activity, and Ca^{2+} is important in maintaining stability. Sequestering agents such as EDTA inhibit the enzymes. The pH optimum is close to neutral, and stability in general is not as good as in the serine proteases. Metalloproteases are produced by several *Bacillus* species, such as *B. amyloliquefaciens*, *B. cereus*, and *B. polymyxa*. *B. thermoproteolyticus* produces

a remarkable protease (Thermolysin) with a high thermostability (up to 80°C). This enzyme has been the subject of detailed investigations, and the amino acid sequence has been established.

a. Production Methods: The metalloprotease of *B. amyloliquefaciens* is formed together with α-amylase and alkaline protease in the normal fermentation process used for production of these enzymes (Keay *et al.*, 1972). NRRL B-3411 is a suitable strain. To prepare the enzyme without interfering enzymes, mutants free from alkaline protease may be used (Murray and Prince, 1970), but methods for fractionation in good yields have also been developed (Keay *et al.*, 1972). Since the metalloprotease is unstable and may be destroyed during fermentation, a short fermentation time is necessary to obtain maximum yields of this enzyme. Thermolysin is prepared in a high-temperature (55°C) fermentation with a fermentation time of about 24 hr (Endo, 1962).

b. Application: Metalloproteases from several *Bacillus* species have been developed industrially for use as rennet substitutes (Murray and Prince, 1970), but the application has been abandoned as it was unsuccessful due to excessive casein hydrolysis. At present metalloproteases are used for bating in tanneries, to hydrolyze barley proteins in breweries, and for proteolysis in the food industry.

C. Acid Proteases

All microbial acid proteases are of fungal origin. Aspartic acid is present in the active center, and there is a considerable homology with the acid proteases of animal origin.

1. Mucor Proteases

As early as 1921 a milk-coagulating protease was made from *Mucor rouxii* (Kohman *et al.*, 1927). It was not successful because the milk-coagulating activity relative to the proteolytic activity was too low. Since then numerous attempts have been made to find a good milk-coagulating enzyme. In a large screening program performed in the early 1960s, Arima and his co-workers (Arima *et al.*, 1968) succeeded in finding a strain of thermophilic *Mucor* which produced a satisfactory enzyme. The strain belonged to the species *M. pusillus*, and the enzyme formed had a higher milk-coagulation-to-protease ratio than any other microbial protease known. The enzyme was subsequently found to give excellent results in actual cheese-making trials. In 1965, it was discovered (Aunstrup, 1968) that the related species *M. miehei* forms a similar enzyme of equally good milk-coagulating properties. Both enzymes are now used as milk coagulants.

M. pusillus and *M. miehei* are thermophiles. The growth temperature range is 20–55°C for *M. pusillus* and 30–60°C for *M. miehei*. A number of morphological and biochemical features distinguish the two species (Cooney and Emerson, 1964). The fact that the two proteases are distinctly different in composition and action, although they belong to the same group of proteases, is of particular interest. It is also interesting that only the *M. miehei* enzyme is a glucoprotein and it is more heat-stable than the *M. pusillus* enzyme.

a. Production Methods: Despite their similarities, the two species require widely different production methods.

M. pusillus is cultivated on a semisolid medium consisting of 60% wheat bran and water for 3 days at 30°C. The enzyme is then extracted with water. The yield is approximately 3200 Soxhlet units per gram of wheat bran (Arima, 1964). Ammonium salts added to the bran will improve the yield.

M. miehei is grown in submerged culture, e.g., in a medium containing 4% starch, 3% soybean meal, 10% ground barley, and 0.5% $CaCO_3$ for 7 days at 30°C. The yield reported corresponds to approximately 3500 Soxhlet units per milliliter of broth (Aunstrup, 1968). A suitable strain is CBS 370.65.

During cultivation the organisms will secrete other enzymes such as lipase, esterase, amylase, and cellulase. If autolysis of the mycelium occurs during fermentation, intracellular protease will leak out. This enzyme has a low milk-coagulating activity, and its presence in the preparation should be avoided. The amylase is only present in small amounts and it is of no significance to the application. Lipase and esterase are usually undesirable in the final preparation.

In the recovery process, the mycelium is first removed by filtration and a concentrated liquid product is prepared by vacuum evaporation or reverse osmosis. Finally, sodium chloride in a concentration of about 20% is added as a preservative. To refine the process, a step for removal of the lipase is usually introduced. This may be done by treatment at pH values between 2 and 3.5 for a short time. If nonspecific protease occurs in the broth, it may be removed by adsorption to silicium dioxide at pH 5. (Moelker and Mattijsen, 1967).

Usually these enzymes are marketed in liquid form; however, solid preparations are used in some areas. They are made by solvent precipitation or direct spray-drying of the purified broth.

b. Application: The commercial preparations are sold in concentrations ranging from 10,000 to 150,000 Soxhlet units. Since the specific activity of the enzyme is approximately 5 million Soxhlet units, the concentration of active enzyme protein in the preparation is about 0.2–3%.

When *Mucor* protease is applied to coagulate milk, a number of parameters influence the activity in a way different from that of calf rennet. There is also a difference between enzymes from *M. pusillus* and *M. miehei*. The variation in temperature, Ca^{2+} concentration and pH are of special importance. By adjusting the reaction conditions during milk coagulation, it is possible to produce cheese of a quality similar to cheese made with calf rennet with both *Mucor* proteases.

2. *Protease from Endothia parasitica*

This enzyme was discovered in 1963 by Sardinas. The species, which is pathogenic to chestnut trees, is apparently the only one of the genus that forms this type of protease (Sardinas, 1966).

a. Production Methods: A suitable strain, e.g., ATCC 14.729, is grown in submerged culture in a medium composed of 3% soybean meal, 1% glucose, 0.3% $NaNO_3$, 1% skim milk, 0.05% KH_2PO_4, and 0.025% $MgSO_4 \cdot 7H_2O$. Fermentation takes place at 28°C and pH 6–7. After 48 hr an enzyme activity of about 2000 Soxhlet units is obtained.

Due to the instability of the enzyme, recovery must take place quickly, at low temperature, and preferably without oxygen. After the mycelium is removed by filtration, the broth is concentrated and the enzyme precipitated by solvent. This enzyme is always marketed as a solid preparation.

b. Application: The milk-coagulating activity is less dependent on pH variation in milk than calf rennet, and the dependence on variation in Ca^{2+} concentration follows that of calf rennet. The thermostability is low; the enzyme is inactivated in less than 5 min at 60°C. It has found limited use as milk coagulant mainly because of its high proteolytic activity. Only in the production of Emmenthal cheese has it proved superior to the *Mucor* proteases. The reason is that this cheese undergoes a high-temperature treatment where the protease is rapidly inactivated. There have been no reports on the presence of lipase or other undesirable enzyme activities in this preparation.

3. *Aspergillus Proteases*

The most important application of *Aspergillus* proteases is for production of soy sauce, of which consumption in Japan alone is over 10^9 liters yearly. In this process bran cultures of *A. oryzae* and the closely related *Aspergillus sojae* are used to hydrolyze the soybean proteins almost completely to amino acids. The extensive hydrolysis is possible because these fungi produce several proteinases and exopeptidases.

Commercial protease preparations made from *A. oryzae* contain acid, neutral, and alkaline proteases. Several components of each type

have been isolated. As a result, the preparations show proteolytic activity in the pH range 4–11. Commercial proteases are also made from black *Aspergillus* species, primarily *A. phoenicis* (syn. *A. saitoi*) and *A. niger* var. *Macrosporus*. These organisms produce acid proteases only, and they are active in the pH range 2–6.

a. Method of Production: The *Aspergillus* proteases can be produced in high yield only by semisolid cultivation. Wheat bran or rice bran is used, and it has been reported that a high ratio of inorganic nitrogen to carbon in the medium is important to obtain good yields (Yoshida and Ichishima, 1964).

A. oryzae NRRL 2160, *A. saitoi* ATCC 14.332, and *A. niger* ATCC 16.513 are useful strains for production of protease. During cultivation several other enzymes are formed in high concentrations: α-amylase, glucoamylase, cellulase, and pectinase. Usually, all these enzymes will be present in the final product unless it is highly purified.

The usual recovery method involves extraction of the bran with water and precipitation of the extract with solvent. More refined, laboratory purification methods are used to make the highly purified preparations used for certain medical purposes. The *Aspergillus* proteases are marketed in solid form.

b. Application: Aspergillus proteases are primarily used in digestive aids, where the acid pH-optimum and the large number of concomitant enzyme activities are beneficial. The market for such preparations in Japan was 520 million yen in 1972 (Samejima, 1974).

In the United States *A. oryzae* is used extensively for flour treatment. The protease exerts a gentle hydrolysis of the gluten, resulting in a considerably reduced mixing time of the dough and an improved bread quality. It is important that the protease be inactivated at a low temperature in the baking process so that extensive degradation of the gluten may be avoided.

Apart from flour treatment, *Aspergillus* proteases find limited application. They may be used in hydrolysis of meat or fish protein under acid conditions to aid the separation from fats, but for economic reasons this application is limited.

VII. AMYLOLYTIC ENZYMES

Starch is an important raw material for food, alcoholic beverages, and various other products. Several enzymatic methods are used to process starch, ranging from slight hydrolysis to reduce the viscosity of starch gels to complete hydrolysis to dextrose. Some of the most important processes and enzymes are summarized in Table IX.

TABLE IX
EXAMPLES OF ENZYMATIC PROCESSES IN THE STARCH INDUSTRY

Process	Enzyme	Source	Reaction conditions		Product
			pH	Temperature (°C)	
Liquefaction	α-Amylase	*B. amyloliquefaciens*	5.5–7	90	Maltodextrins
		B. licheniformis	5.5–9	110	DE 10–20[a]
Debranching	Pullulanase	*K. pneumoniae*	6–7	50–60	Intermediate process in the manufacture of dextrose
Saccharification	α-Amylase	*A. oryzae*	5–7	50–55	High-maltose syrup High-DE syrup
Saccharification	Glucoamylase	*A. niger* (*Rhizopus* sp.)	4–5	55–60	High-DE syrup Crystalline dextrose
Isomerization	Glucose isomerase	(*Streptomyces* sp.) *B. coagulans* *Actinoplanes* sp.	6.5–8.5	60–65	Fructose syrup

[a] DE = dextrose equivalent.

The α-amylases catalyze a random hydrolysis of the α-1,4-bonds of the starch molecule. The result is a rapid decrease of viscosity and formation of dextrins and small amounts of glucose and maltose. Glucoamylase hydrolyzes from the nonreducing end of starch and dextrin molecules and splits off one glucose molecule at a time. α-1,6-Bonds are attacked but at a slower rate than the α-1,4-bonds. Pullulanase and isoamylase attack α-1,6-bonds of the starch or dextrin molecules.

A. Bacterial α-Amylase

The α-amylase from *B. amyloliquefaciens* was developed for industrial use by Boidin and Effront in 1913 (Boiden and Effront, 1917). The nomenclature of the producing organism has been somewhat uncertain. Initially it was called *B. mesentericus*, then *B. subtilis*, and now *B. amyloliquefaciens*. It is worth mentioning that the commercial amylase is different from the saccharifying amylase produced by the typical *B. subtilis*, the Marburg strain.

The α-amylase is quite heat-stable and may be used in starch hydrolysis up to 90°C. However, the starch industry was interested in a still higher hydrolysis temperature. That was made possible in 1972 by the introduction of an α-amylase made by *B. licheniformis*. This enzyme allows starch hydrolysis up to 105–110°C (Madsen *et al.*, 1973).

The *Bacillus* amylases are metalloenzymes. The *B. amyloliquefaciens* amylase consists of four subunits bound together by one zinc atom. The subunits are separable, and enzyme molecules with multiple subunits may be formed, but the four-subunit structure has the highest activity (Robyt, 1973).

Calcium ion stabilizes the enzymes and is customarily added to the reaction liquids. Amylase from *B. licheniformis* is less dependent upon Ca^{2+} stabilization as is the enzyme from *B. amyloliquefaciens*.

a. Production Methods: The *Bacillus* amylases are always prepared in submerged culture. The medium used is relatively rich and may contain ground grain such as corn and a protein source such as soybean meal or corn steep liquor. Additional carbohydrates may be added during the fermentation but this must be done carefully to avoid overdosage, as glucose represses the amylase formation.

Fermentation usually takes place at a temperature between 30° and 40°C, and pH is around neutral. If pH falls below 6, the amylase activity will be lost because of denaturation. In order to prevent this and to improve stability of the amylase, $CaCO_3$ may be added to the medium. The α-amylase formation starts as soon as the bacterial count approaches its maximum of 10^9 to 10^{10} cells/ml after about 10–20 hr. It

continues until the utilizable carbon source is exhausted—usually after 100–150 hr. It is interesting that *B. amyloliquefaciens* will utilize maltose and the lower saccharides up to a pentasaccharide at a very slow rate, whereas *B. licheniformis* utilizes these carbohydrates much faster. Apparently, these carbohydrates are not hydrolyzed by the α-amylase but have to be hydrolyzed by an α-glucosidase before they can be utilized. Both microorganisms possess this enzyme, but *B. amyloliquefaciens* has much less than *B. licheniformis*. The presence of this enzyme may also explain the transglucosylation reactions which have been observed in some crude α-amylase preparations. The α-glucosidases are unstable and are formed in small amounts; consequently, they are not present in the commercial products.

Both organisms form serine protease during cultivation. *B. licheniformis* does not form other extracellular enzymes in significant amounts, but *B. amyloliquefaciens* will form a neutral protease, β-glucanase, and hemicellulase. Of these, only the proteases are undesirable in starch hydrolysis because they catalyze the formation of melanin products.

Consequently, most commercial α-amylase products are now protease-free. This condition can be obtained in various ways. The simplest method is to heat the broth to a temperature at which the protease activity is destroyed but only a minimum of the amylase is lost. Another possibility is separation of the protease by adsorption to clay minerals (e.g., bentonite) or fractional precipitation. The advantage of the latter method is that the protease can be recovered. However, it is difficult to remove the protease activity completely in this way. The method is quite expensive because of the chemicals required and because some α-amylase activity is lost in the process. The most sophisticated solution to the problems is to use a mutant incapable of producing the proteases. Such mutants may cause problems in the preparation of the medium as it is necessary to use prehydrolyzed protein. The fermentation yield is decisive in the economy of the production process; high-yielding strains, therefore, are important. Outtrup reported a 25-fold increase in productivity of a *B. licheniformis* strain (Outtrup and Aunstrup, 1975) in six mutational steps. Bacterial α-amylases are usually marketed as liquid preparations which are preserved with 20% sodium chloride.

Activities of commercial preparations vary by a factor of 100 and are adjusted to the application in question. There is a tendency to produce more concentrated products because of freight and packing savings. The most active liquid preparations may contain 2% active amylase protein, and the most active solid preparations 5% active amylase protein.

b. Properties and Application: B. licheniformis amylase has a wider pH range than the *B. amyloliquefaciens* amylase. It is much more heat-stable and less dependent on Ca^{2+}. This enzyme has only two disadvantages: it is very difficult to inactivate by heat treatment, and it is more expensive if activities are compared at low temperatures. At the application temperature the *B. licheniformis* amylase, in fact, costs less than the *B. amyloliquefaciens* amylase.

The enzymes are not stable at pH values below 6, and their optimum activity is ~6.5–7. This means that the pH must be raised to this level in many applications (e.g., treatment of raw starch) where initial pH is low.

Bacterial α-amylases have two large fields of application. The earliest is desizing, for which amylase is used to remove starch sizes from the warp of cotton fabrics. The most important use is starch liquefaction in preparation of glucose syrups, dextrose, or fructose/glucose syrups. In both applications the *B. licheniformis* enzyme has proved superior to the *B. amyloliquefaciens* enzymes due to better heat stability.

Liquefaction of starchy raw materials for production of alcohol is another important application. Bacterial α-amylases have almost completely superseded the traditional malt in this process. The brewing industry also employs α-amylase with advantage in the liquefaction of various starch raw materials that are used as adjuncts.

B. Fungal α-Amylase

α-Amylase is the main component of the old enzyme preparation Taka-diastase prepared from *A. oryzae*. The enzyme is called Taka-amylase A by some authors. The molecule is different from the *Bacillus* amylase in a number of points: no subunits have been detected; it contains eight half-cystine groups and one SH group; it is a glycoprotein with 8 moles mannose, 1 mole xylose, and 2 moles hexosamine. Ten calcium ions are associated with the molecule, nine of which may be removed by dialysis. The molecular weight is ~51,000. *A. oryzae* amylase has pH optimum at 4.8–5.8, and it is less heat-stable than the *B. amyloliquefaciens* amylase.

α-Amylase is also produced by *A. niger*. The properties of this enzyme are similar to those of the *A. oryzae* enzymes, but some *A. niger* strains produce an additional acid amylase which is fairly stable down to pH 2 and somewhat more heat-stable. Despite the obvious practical advantages of this enzyme, it has found only limited application, probably because of low yield and, consequently, high price.

a. Production Method: The traditional production method is cultivation of the fungus on wheat bran. A number of enzymes apart from the

α-amylase are produced this way, but by adjusted medium composition and a selected strain it is possible to optimize the α-amylase production. Products made this way usually contain fairly high concentrations of proteinase. *A. oryzae* amylase may also be produced by cultivating a strain in submerged culture in media similar to those used to make *Bacillus* amylases. It is characteristic of this fermentation that the viscosity created by the mycelium is high, so that aeration and stirring become a problem. Glucose inhibits the amylase formation, and addition of glucose during fermentation must therefore be done with care. Amylase preparations made by submerged fermentation contain fewer other enzymes than preparations made in semisolid cultivation. Primarily, the protease content is much lower. Both preparations have little glucoamylase activity.

b. Properties and Applications: Aspergillus amylase is more saccharifying, i.e., produces more sugars, than do the equivalent *Bacillus* enzymes. It is possible to obtain over 50% of maltose when starch is hydrolyzed by this enzyme. Syrups of high maltose concentration are useful for a number of purposes where special functional properties are required.

Another important application of *A. oryzae* amylase is in the baking industry, where the enzyme is added to flour with a low amylase content. The amylase degrades starch in the dough, and the maltose formed serves as substrate for the baker's yeast during leavening. The low temperature stability of the *A. oryzae* amylase is important as it prevents extensive degradation of the crumb during baking.

C. Glucoamylase

Glucoamylases have been found in several genera of fungi, but only organisms belonging to *Aspergillus*, *Rhizopus*, and *Endomyces* have been used for commercial production. All three types of glucoamylases are capable of completely hydrolyzing starch to glucose. Enzymes from *Rhizopus* and *Endomyces* are somewhat less thermostable than the *Aspergillus* enzyme. The optimum temperature for starch hydrolysis is 60°C for the *Aspergillus* enzyme and 55°C for the other two enzymes. The difference is important because it is difficult to prevent microbial contamination in the hydrolysis process at temperatures below 60°C. Consequently, the *Aspergillus* enzyme is the preferred glucoamylase. The other two enzymes find limited use, primarily in the Soviet Union and Japan.

Glucoamylase is produced by strains of the *A. niger* group. Patent literature describes processes based on several species belonging to this group, e.g., *A. niger*, *A. awamori*, *A. phoenicis*, *A. diastaticus*, and

A. usamii. No comparative investigations have been published on the enzymes from these organisms, but the differences appear to be as small as the differences in morphology between the species. The following, therefore, describes the enzyme as *A. niger* glucoamylase.

The kinetic and molecular properties of *A. niger* glucoamylase have been studied in detail (Freedberg *et al.*, 1975), but a general agreement on the results has not yet been reached. At least two enzyme components are found in the cultured broth. They have similar activities except for the ability to attack raw starch and glycogen. The molecular weight is 60,000–100,000, and the molecules contain 13–18% carbohydrate. The optimum pH is 4.2–4.5, and the enzyme catalyzes the hydrolysis of α-1,4- and α-1,6-glucosidic bonds at a decreasing rate as the chain length decreases. The temperature stability is good up to 60°C at optimum pH.

a. Production Methods: The *Aspergillus* strain is grown in submerged culture in a rich medium containing a high concentration of starch. A good composition is 20% corn and 2.5% corn steep liquor (Smith and Frankiewies, 1975). The starch must be liquefied by a heat-stable bacterial α-amylase before sterilization. During the hydrolysis and in the initial stage of fermentation a large concentration of glucose is present in the medium. Unlike most other hydrolytic enzymes, formation of glucoamylase is not repressed by glucose. During fermentation pH drops to 3–4; it may be controlled by NaOH or by introduction of NH_3 into the aeration system (Dworschack and Nelson, 1972). Fermentation takes place at 30–35°C for about 4–5 days. During this time the starch is completely hydrolyzed, and the glucose formed is metabolized by the fungus. The metabolic products are mainly CO_2, but a significant amount remains in the broth in the form of organic acids and sugar alcohols. Minor side activities such as protease, cellulase, lactase, α-amylase, and transglucosidase are formed during fermentation.

The α-amylase activity will be denatured quickly in the broth if pH is below 4.5, and small amounts are of no importance for the application. The transglucosidase is undesirable in the application and must removed if present. Good strains which produce very little transglucosidase are known (e.g., *A. awamori* NRRL 3112) (Smiley, 1967). With most strains the removal of transglucosidase is a considerable problem in the recovery process, and many methods for its removal have been patented (Table VIII). It is not known to what extent these methods have been used commercially. Several of the methods are unattractive because of loss of glucoamylase activity in the process.

Glucoamylase is almost exclusively marketed in liquid form. The

microbial stability is good owing to the low pH (~4.5) and the high dry-substance concentration caused by the presence of sugar alcohols and organic acids. The sugar alcohols (mannitol) also have a stabilizing effect on the activity when the enzyme is stored. Solid preparations may be made in the usual way, but it is often necessary to reduce the amounts of sugar alcohols and organic acids first, for instance, by ultrafiltration.

b. Properties and Application: The content of active enzyme protein in commercial products is usually high, up to ~5%. The products contain small amounts of other enzyme activities such as protease, α-amylase, and cellulase.

The primary application of glucoamylase is in starch saccharification, which is described in detail by Antrim *et al.* (this volume). Glucoamylase has another useful application, viz., production of alcohol from starchy materials. Glucoamylase from *Rhizopus* was used for this purpose more than 70 years ago in the so-called amyloprocess. It was not very successful, presumably because the technology required for the process was too advanced for its time. In 1947, a new technique was developed at the Northern Regional Research Laboratory, Peoria, Illinois. *A. niger* was the enzyme source, and the process is still used in some large distilleries. The method is similar to that normally used in the production of glucoamylase, but all of the broth is transferred to the mash to be saccharified.

In many distilleries commercial glucoamylase is used extensively, combined with bacterial α-amylase, for liquefaction of the starch. Claims of improved alcohol yield relative to malt have been made and may be due to the presence of enzymes capable of hydrolyzing glucosides other than starch, for instance gums or cellulose. Glucoamylase finds minor application in the manufacture of low-calorie beer. It is added to the fermentation vat and hydrolyzes the dextrins to glucose, which is subsequently fermented by the yeast.

VIII. OTHER ENZYMES

A. Pectinase

This term covers at least six different enzymes which take part in the hydrolysis of the pectin molecule. Commercial products contain a mixture of these enzymes adapted to the application in question. The enzymes are produced by cultivation of *A. niger* or *A. wentii* in submerged or semisolid culture. Pectin may be added to the medium as an inducer.

The enzyme composition depends on the method of cultivation. In submerged culture, polygalacturonase is the predominant enzyme, whereas pectinesterase is scarce. Semisolid culture produces a mixture of enzymes that is more satisfactory for most applications.

Pectinases are used in preparation of fruit juice and wine. By means of pectinase the yield may be improved, viscosity adjusted, or special effects such as clarification obtained. It is possible, for instance, to improve the stability of cloudy orange drinks or to make clear lemon drinks by application of special pectinase preparations. Application of pectinases is a question of experience; it is made difficult by seasonal and geographic variations of the composition of the fruit that has to be treated. Furthermore, the complexity of the enzyme preparation makes exact standardization of the various pectinases extremely difficult. Extensive laboratory tests are required for each new crop of fruit in order to determine the proper pectinase treatment.

B. β-Glucanase

Commercial β-glucanase preparations are used to hydrolyze barley gums in the brewing industry. Some 20 years ago a considerable market was expected in the chicken feed industry, but it has not yet materialized. The purpose was to improve the feed value of certain types of hard barley by hydrolyzing the gums.

β-Glucanase is produced by *B. amyloliquefaciens* and synthesized concomitant with α-amylase and protease. Commercial preparations are optimized for the glucanase content by adjusting fermentation conditions accordingly. The enzyme is an endoglucanase, and the hydrolysis products are saccharides with 3–5 glucose molecules. The enzyme is stable at pH values above 5 and up to 50°C.

Fungal β-glucanase is made by fungi of the *A. niger* group. It is often formed as a side activity in pectinase preparation. Commercial preparations are optimized for production of β-glucanase. Since the enzyme preparations usually contain several β-glucosidases apart from the β-glucanase, an extensive hydrolysis (to glucose) is possible. The enzyme has a pH-optimum ~5 and may be used in combination with bacterial glucanase in the brewing industry for mashing or during fermentation if desired.

C. Pullulanase

This is used as a debranching enzyme in starch hydrolysis. Several microorganisms produce pullulanase, but the only commercial product available is prepared from *Klebsiella aerogenes* (previously known

as *Aerobacter aerogenes*). The enzyme has a pH-optimum at 5 and is rapidly inactivated at temperatures above 50°C.

Pullulanase is produced in submerged culture in a medium containing hydrolysates of starch high in amylopectin (Bulich, 1976) and in nitrogen sources, such as corn steep liquor. The fermentation takes place between 25 and 35°C, pH is maintained between 7 and 8. A suitable strain is *K. aerogenes* NRRL B 7580. The enzyme is partly cell-bound and the ratio of cell-bound enzyme to free enzyme varies according to the composition of the medium. The ratio is influenced by the carbohydrate component of the medium, and by using high-molecular-weight amylopectin it is possible to obtain over 75% free enzyme. Before recovery the cell-bound enzyme is released from the cells by the use of a nonionic surfactant.

D. Dextranase

Dextranases are used to a limited extent in the sugar industry when dextrans occur in the sugar-containing juice. Commercial preparations are made from *Penicillium funiculosum* or *Penicillium lilacinum*, pH-optimum is ~5, and they may be used up to 50° or 60°C. The enzyme from *P. lilacinum* is preferred because it has the best heat stability.

Dextranases have been advocated as a means of removing dental plaque. Enzymes such as the aforementioned two are of little value, as the dextrans of dental plaque are usually of the α-1,3-linkage type, whereas the penicillium enzymes hydrolyze α-1,6-bonds only. Enzymes suitable for removal of dental plaque have been described in the literature, but no preparations have yet been marketed.

E. Cellulase

Cellulase has enjoyed much publicity because of its potential application in waste treatment. The current market is small; the most important use is for digestive aids.

The best organism for production of cellulase is *Trichoderma reesei*. Enzyme from this organism is preferred because it shows relatively good activity against native cellulose. Most *Aspergillus* enzyme preparations contain cellulolytic activity against cellulose derivatives, such as carboxymethylcellulose, but no activity against native cellulose.

Cellulase may be prepared in semisolid or submerged culture. The enzyme may be adsorbed to cellulose present in the medium. Sophorose is said to be an inducer of the enzyme. Most cellulase preparations are solid; this only reflects the state of the art. If a bigger field of application is discovered, liquid preparations will probably

soon be developed. The products are relatively expensive, and the high cost is one of the elements that have prevented general use.

F. Lactase (β-Galactosidase)

This enzyme is intracellular in bacteria and yeasts, but it is extracellular in many fungi, for instance, *Aspergilli*, *Mucor* sp., *Rhizopus* sp., and *Penicillium* sp. Commercial preparations are made from *A. oryzae* and *A. niger*. They have acid pH-optima and are consequently well suited for use in digestive aids. Lactase from *A. oryzae* has pH-optimum of 4.8, and will have about 10% of its maximum activity at pH 2. Heating to 60°C for 10 min causes an 85% loss of activity. The *A. niger* lactase is more resistant; the pH-optimum is 3.5, 50% of maximum activity is at pH 2, and it will resist heating for 1 hr at 55°C without significant loss of activity.

The *Aspergillus* lactases are produced by semisolid cultivation in acidified wheat bran at 30°C. *A. oryzae* ATCC 20423 (Kiuchi, 1975) and *A. niger* ATTCC 13496 (Cayle, 1971) are suitable strains. The enzyme is extracted from the bran with water and may be precipitated directly with solvent or purified by adsorption to bentonite at pH 4, washing and releasing the enzyme at pH 7; it is then recovered by solvent precipitation. These enzymes are usually of high purity and therefore rather expensive. They are used mainly for digestive aids.

G. Lipase

Lipases are formed by many microorganisms. The traditional sources for commercial products are *Rhizopus* sp., *Mucor* sp., *Aspergillus* sp., and *Candida* sp. Some of the microorganisms are known to produce several lipases, and the specificity of the lipases varies both with regard to fatty acids and position in the triglyceride molecule.

The method preferred for preparation is semisolid fermentation, and the enzymes are recovered in the usual way. Lipids are not necessary for induction of the enzyme production and may in some instances inhibit enzyme synthesis. Lipase may be recovered as a by-product in the production of microbial rennet by *M. miehei* by absorption on clay minerals at pH 5 and elution at pH 10 (Moskowitz *et al.*, 1975). *Candida* lipases are usually prepared in submerged culture. Hydrocarbons have been reported as carbon sources, but not in commercial production.

Numerous applications have been proposed (Seitz, 1974), but their use has been limited. The most important market is digestive aids, for which microbial lipases replace pancreatic lipase, which is expensive and scarce. It is also used to improve the flavor of special types of

cheese. Lipases have not been introduced in detergents, although extensive investigations have been made and several patents issued on the use of lipases for this purpose.

IX. CONCLUSION

The extracellular enzymes are the "work horses" of the enzymologist, not very sophisticated but tough and hard-working. In many cases their use is hampered by high prices or properties which do not fit the process so well. Future developments will probably be directed to alleviation of these problems: yield improvements to decrease costs and screening for new organisms, modification of structure genes, or derivation of enzyme molecules by chemical means to improve enzyme properties.

No matter to what extent these attempts are successful, there is little doubt that the hydrolytic activities of the extracellular enzymes will maintain their position as important industrial tools. Their specificity, mild reaction conditions, and low toxicity are properties that will all become increasingly significant.

In regard to the economic question, which is conclusive in most business decisions, enzymes have the advantage of high specific activity. The actual raw-material costs are therefore insignificant, and the costs of preparation are open to attack by genetic and biochemical engineering—a real challenge.

REFERENCES

Arima, K. (1964). U.S. Patent No. 3,151,039.

Arima, K., Yu, J., Iwasaka, S., and Tamura, G. (1968). *Appl. Microbiol.* **16**, 1727.

Aunstrup, K. (1968). British Patent No. 1,108,287.

Aunstrup, K. (1977). *In* "Biotechnology and Fungal Differentiation" (J. Meyrath and J. D. Bu'Lock, eds.), p. 157. Academic Press, New York.

Aunstrup, K., and Outtrup, H. (1973). British Patent No. 1,303,633.

Aunstrup, K., Outtrup, H., Andresen, O., and Dambman, C. (1972). *Proc. Int. Ferment. Symp., 4th, Tokyo*, p. 229.

Aunstrup, K., Andresen, O., and Outtrup, H. (1973). U.S. Patent No. 3,723,250.

Boidin, A., and Effront, J. (1917). U.S. Patent No. 1,227,374.

Bulich, A. A. (1976). U.S. Patent No. 3,963,575.

Cayle, T. (1971). U.S. Patent No. 3,629,073.

Churchill, B. W., Steel, D. R., and Buss, D. R. (1973). U.S. Patent No. 3,740,318.

Cooney, D. G., and Emerson, R. (1964). "Thermophilic Fungi." Freeman, London.

Dwonschack, R. G., and Nelson, C. A. (1972). U.S. Patent No. 3,660,236.

Endo, S. (1962). *J. Ferment. Technol.* **40**, 346.

Feldman, L. I. (1971). U.S. Patent No. 3,626,957.

Freedberg, I. M., Levin, Y., Kay, C. M., McCubbin, W. D. and Katchalsky-Katzir, E. (1975). *Biochim. Biophys. Acta* **391**, 361.
Güntelberg, A. V. (1954). *C.R. Lab. Carlsberg, Ser. Chim.* **29**, 27.
Hoerle, R. D. (1976). Danish Patent Application No. 1618/77.
Horikoshi, K., and Ikeda, Y. (1977). U.S. Patent No. 4,052,262.
Jensen, D. E. (1972). *Biotechnol. Bioeng.* **14**, 647.
Keay, L., and Anderson, R. G. (1971). U.S. Patent No. 3,592,737.
Keay, L., Moseley, M. H., Anderson, R. G., O'Connor, R. J., and Wildi, B. S. (1972). *Biotechnol. Bioeng. Symp.* No 3, 63.
Kiuchi, A., and Tanaka, Y. (1975). U.S. Patent No. 3,919,049.
Kohman, H. A., Irwin, R., and Stateler, E. S. (1927). U.S. Patent No. 1,654,176.
Lehmann, R., Pfeiffer, H. F., Schindler, J., and Schreiber, W. (1977). U.S. Patent No. 4,062,732.
Madsen, G. B., Norman, B. E., and Slott, S. (1973). *Staerke* **25**, 304.
Moelker, H. C. T., and Mattijsen, R. (1967). French Patent No. 1,592,965.
Moskowitz, G. J., Como, J. J., and Feldman, L. I. (1975). U.S. Patent No. 3,899, 395.
Murray, E. D., and Prince M. P. (1970). U.S. Patent No. 3,507,750.
Outtrup, H., and Aunstrup, K. (1975). *Proc. Int. Congr. IAMS, 1st*, **5**, 204.
Outtrup, H., Andresen, O., and Aunstrup, K. (1972). British Patent No. 1,296,839.
Priest, F. G. (1977). *Bacteriol. Rev.* **41**, 711.
Robyt, J. F., and Ackerman, R. J. (1973). *Arch. Biochem. Biophys.* **155**, 145.
Samejima, H. (1974). *Enzyme Eng.* **2**, 363.
Sardinas, J. L. (1966). U.S. Patent No. 3,275,453.
Schleich, H. (1971). U.S. Patent No. 3,616,233.
Seitz, E. W. (1974). *J. Am. Oil Chem. Soc.* **51**, 12.
Smiley, K. (1967). U.S. Patent No. 3,301,768.
Smith, J. A., and Frankiewies, J. R. (1975). German Patent Aus. No. 25-54-850.
Smythe, C. V., and Drake, B. B. (1949). U.S. Patent No. 2,480,090.
Smythe, C. V., Drake, B. B., and Neubeek, C. E. (1950). U.S. Patent No. 2,530,210.
Welker, N. E., and Campbell, L. L. (1967). *J. Bacteriol.* **94**, 1124.
Yamada, N., and Tomoda, K. (1966). U.S. Patent No. 3,293,114.
Yoshida, F., and Ichishima, E. (1964). U.S. Patent No. 3,149,051.

Journal of Food Protection, Vol. 46, No. 5, Pages 453-468 (May 1983)

Determining the Safety of Enzymes Used in Food Processing

M. W. PARIZA and E. M. FOSTER

Food Research Institute, Department of Food Microbiology and Toxicology, University of Wisconsin, Madison, Wisconsin 53706

(Received for publication October 18, 1982)

ABSTRACT

Enzymes are proteins that catalyze chemical reactions. They are highly specific and needed in only minute quantities. Certain enzymes have long been used to produce specific foods (e.g., cheese). Today they have numerous applications and are increasing in commercial importance. There has never been a health problem traced to the use of an enzyme per se in food processing. However, it is important that scientific data be provided to show that enzyme preparations, particularly those lacking a long history of safe use, are in fact safe to consume. The purpose of this report is to propose guidelines for assessing enzyme safety. We conclude that the enzymes per se now used or likely to be used in the future in food processing are inherently nontoxic. Safety evaluation should focus on possible contaminants which could be present. Assuming that current Good Manufacturing Practices (CGMPs) are followed, toxic contaminants could only come from the enzyme source itself (animal, plant or microbial). Hence, the safety of the source organism should be the prime consideration. Enzymes from animals or plants commonly regarded as food need not be subjected to animal feeding studies. Some food plants produce toxins and chemical assays may be used in these cases to assess safety. For enzymes from bacteria, it should be shown that antibiotics and acute toxins active via the oral route (enterotoxins and certain neurotoxins) are absent. Small molecular weight toxins (< 500 daltons) may be produced by certain fungi and actinomycetes. It should be shown that enzymes from such organisms are free of these materials. If it is established that a microbial culture does not produce antibiotics or toxins active via the oral route, then enzymes manufactured from that culture using CGMPs may be regarded as safe for use in food processing.

BACKGROUND

To understand and apply the proposed guidelines for determining safety of enzymes used in food processing, it is necessary to consider what enzymes are, how they act, how they are prepared and how they are used. That is the purpose of this section.

General considerations

Enzymes are proteins which catalyze chemical reactions. Like all catalysts enzymes increase the rates at which reactions achieve equilibrium. For example, there are instances where certain enzymes increase the rates of specific reactions by 10 million times (*47*). Enzymes act by lowering activation energy. Since they cannot create energy, enzymes will only affect reactions which, because of a "downhill" net energy flow, could occur spontaneously. Like other catalysts, enzymes are not consumed by the reactions which they catalyze. Hence, one enzyme molecule can, through time, catalyze the transformation of many molecules of substrate (*47, 52*).

Most complex chemical reactions not controlled by catalysts produce a variety of products. However, in general, enzymes accelerate specific reactions which result in the generation of specific products. High degrees of specificity and strong catalytic activities are the most important functional properties of enzymes. Clearly, without enzymes DNA could not be replicated nor could RNA and proteins be synthesized and degraded. The controlled and orderly array of metabolic processes of living cells, which in fact define life, would not be possible. Life on earth is absolutely dependent upon enzymes. Every cell comprising every organism alive at this moment contains enzymes which are functioning in highly ordered and specific ways to transform one chemical into another as dictated by biological necessity.

Like all proteins, enzymes are synthesized inside cells by a complex process involving DNA, RNA, cellular structures called ribosomes, various small molecules such as amino acids, energy-rich phosphorus compounds and certain cations, and enzymes to catalyze specific reactions (*52*). The fact that enzymes are a necessary component in the biological mechanism which produces new enzymes underscores the fundamental importance of these remarkable biological catalysts.

After synthesis, enzymes may remain inside cells or they may be secreted into the extracellular milieu. Secreted enzymes are hydrolytic and their purpose is to decompose macromolecules into small units which then can be taken up by cells and used (under enzymic direction) as needed in metabolic processes. Enzymes which remain inside cells (intracellular) are of all classes and may be involved in synthesis or degradation of various substances. Economically important enzymes are found among both the intracellular and extracellular groups (*47*).

The name given to an enzyme is determined according to the reactions which is catalyzed. It is customary to attach the suffix "-ase" to the name of the principal sub-

strate upon which the enzyme acts; e.g., the sugar *lactose* is acted upon by *lactase*, *proteins* are degraded by *proteases*, intramolecular rearrangements *(isomerizations)*are catalyzed by *isomerases*. Additionally, many well-known and long-used enzymes have trivial (common, historical) names, e.g., papain from papaya. To minimize confusion, each enzyme activity is assigned a four-part number (called the IUB[1] number) and a systematic name based on the reaction. However, this system does not distinguish between different enzymes from different organisms which catalyze the same reaction (*47*).

All living organisms produce and contain many enzymes, but no one organism has enzymes for all or even most possible biotransformations. Organisms may produce one specific enzyme to act on a given substrate. Organisms may also produce two or more different enzymes which catalyze the same reaction; such enzymes are called isoenzymes. The reasons for this are not known, but it is believed related to the apparent necessity of organisms to maintain precise control over enzyme synthesis, degradation and activity (*52*). Although enzymes catalyzing the same reaction but produced by different species may be similar, it is also possible that they may be entirely different (*21, 52*). Similarities and differences between enzymes and other proteins is one way of estimating evolutionary divergence among species (*21, 52*).

Catalytic activity is ultimately derived from the sequence of specific amino acids which comprise an enzyme. Amino acid sequence, in turn, determines the shape of the enzyme molecule. The shape or configuration is all-important. Disrupting the shape destroys activity.

Enzyme activity is operationally defined by kinetic parameters such as maximum catalytic rate and the affinity of the enzyme for its substrate. Virtually any environmental factor (pH, ionic strength, temperature, etc.) affects enzyme activity. Enzymes are also subject to inhibition by various means (*47, 52*). These properties permit cells to regulate the activities of enzymes which they synthesize and contain. A thorough understanding of the properties of individual enzymes also permits their optimal use in industry.

Historical examples of enzyme use

Most of what we call "food" is really tissue derived from living organisms (animals or plants); in some cases (e.g., milk), food is a secretion from living cells. Many of the enzymes in the cells of tissues remain active after cell death. For example, meat is "aged" by hanging animal carcasses in refrigerated rooms for several days after slaughter. During this time cells in the tissues break down, freeing various degradative enzymes, which then partially digest the connective tissue to give a more tender product. The tenderizing process can be accelerated by adding proteolytic enzymes derived from other sources to the meat at various stages before consumption, such as injecting proteases into the vascular system of the animal before slaughter or sprinkling papain (protease from papaya) on the meat before cooking. The tenderizing process is simply the first step in digestion which continues in the gastrointestinal tract of the consumer.

Enzymes have always been present in human food even though they have only recently been recognized as such. In addition to tissue-derived enzymes, microorganisms (because they are ubiquitous) also pervade the food supply, and the enzymes in microorganisms can alter the character of food. It was discovered early in the development of human civilization that some microbial transformations are desirable.

One of the first to be recognized was the souring of milk, a necessary step in making cheese. According to legend, cheesemaking was discovered several thousand years ago when an Arabian merchant carried milk in a pouch made of sheep's stomach. Rennet in the lining of the pouch caused the milk to curdle. We must assume that microorganisms grew at the same time and produced other enzymic changes that came to be regarded as desirable.

During the intervening centuries, man has learned how to make hundreds of kinds of cheese by controlling the environment and by adding types of microorganisms that produce enzymes which can bring about desirable changes. Lipases and proteases from various animal and microbial sources can also be added to achieve certain desired qualities.

We now use the term "fermentation" to describe milk souring and similar processes involving mass growth of microorganisms to produce useful products (*52*). Originally, however, the term described the transformation of grape juice into wine. Production of wine from grapes through fermentation also has its origin in antiquity. Among the treasures placed in the tombs of Egyptian pharaohs were casks of wine. The ancient Greeks attributed to the god Bacchus the discovery of fermentation (*52*). We now know that it is not yeast per se, but rather a system comprised of several enzymes contained in yeast that is ultimately responsible for the production of ethanol and carbon dioxide from the sugar in grape juice. This enzyme system was one of the first to be extensively studied and characterized. In fact, the word "enzyme", introduced by Kuehne, means "in yeast," although it has been expanded and now applies to all proteinaceous catalysts from any biological source (*52*).

Other ancient processes of food alteration and/or preservation involving enzymic action include breadmaking (yeast) and the production of vinegar from wine (*Acetobacter*). Only within the past 100 years has it been recognized that enzymes exist as discrete entities, and can, in fact, function in isolated systems outside living cells (*52*). This realization has led to remarkable advances through technological application of enzymes to many areas of human need.

Modern uses of enzymes

Food processing. Fermentations involving living or-

[1]The enumeration system of the Enzyme Commission of the Third International Congress of the International Union of Biochemistry (*47*).

TABLE 1. *Enzyme preparations used in food processing (3).*

Trivial name	Classification	Source	Systematic name (IUB)[a]	IUB No.[a]
α-Amylase	Carbohydrase	(1) *Aspergillus niger*, var. (2) *Aspergillus oryzae*, var. (3) *Rhizopus oryzae*, var. (4) *Bacillus subtilis*, var. (5) Barley malt (6) *Bacillus licheniformis*, var.	1,4-α-D-Glucan glucanohydrolase	3.2.1.1
β-Amylase	Carbohydrase	Barley malt	1,4-α-D-Glucan maltohydrolase	3.2.1.2
Bromelain	Protease	Pineapples: *Ananas comosus*, *Ananas bracteatus* (L)	None	3.4.22.4
Catalase	Oxidoreductase	(1) *Aspergillus niger*, var. (2) Bovine liver (3) *Micrococcus lysodeikticus*	Hydrogen peroxide: hydrogen peroxide oxidoreductase	1.11.1.6
Cellulase	Carbohydrase	(1) *Aspergillus niger*, var. (2) *Trichoderma reesei*	1,4-(1,3;1,4)-β-D-Glucan 3(4)-glucanohydrolase	3.2.1.4
Ficin	Protease	Figs: *Ficus* sp.	None	3.4.22.3
β-Glucanase	Carbohydrase	(1) *Aspergillus niger*, var. (2) *Bacillus subtilis*, var.	1,3-(1,3;1,4)-β-D-Glucan 3(4)-glucanohydrolase	3.2.1.6
Glucoamylase (Amyloglucosidase)	Carbohydrase	(1) *Aspergillus niger*, var. (2) *Aspergillus oryzae*, var. (3) *Rhizopus oryzae*, var.	1,4-α-D-Glucan glucohydrolase	3.2.1.3
Glucose isomerase	Isomerase	(1) *Actinoplanes missouriensis* (2) *Bacillus coagulans* (3) *Streptomyces olivaceus* (4) *Streptomyces olivochromogenes* (5) *Streptomyces rubiginosus*	D-Xylose ketolisomerase	5.3.1.5
Glucose oxidase	Oxidoreductase	*Aspergillus niger*, var.	β-D-Glucose: oxygen oxidoreductase	1.1.3.4
Hemicellulase	Carbohydrase	*Aspergillus niger*, var.	None	None
Invertase	Carbohydrase	*Saccharomyces* sp. (*Kluyveromyces*)	β-D-Fructofuranoside fructohydrolase	3.2.1.26
Lactase	Carbohydrase	(1) *Aspergillus niger*, var. (2) *Aspergillus oryzae*, var. (3) *Saccharomyces sp.*	β-D-Galactoside galactohydrolase	3.2.1.23
Lipase	Lipase	(1) Edible forestomach tissue of calves, kids, and lambs (2) Animal pancreatic tissues (3) *Aspergillus oryzae*, var. (4) *Aspergillus niger*, var.	Carboxylic-ester hydrolase Triacylglycerol acylhydrolase	3.1.1.1 3.1.1.3
Papain	Protease	Papaya: *Carica papaya* (L)	None	3.4.22.2
Pectinase[b]	Carbohydrase	(1) *Aspergillus niger*, var. (2) *Rhizopus oryzae*, var.	Poly (1,4-α-D-galacturonide) glycanohydrolase Pectin pectyylhydrolase Poly (1,4-α-D-galacturonide) lyase	3.2.1.15 3.1.1.11 4.2.2.2

Pepsin	Protease	Porcine or other animal stomachs	None	3.4.23.1
Protease (general)	Protease	(1) *Aspergillus niger*, var. (2) *Aspergillus oryzae*, var. (3) *Bacillus subtilis*, var. (4) *Bacillus licheniformis*, var.	None	3.4.21.14 3.4.24.4
Rennet	Protease	(1) Fourth stomach of ruminant animals	None	3.4.23.4
		(2) *Endothia parasitica*		3.4.23.6
		(3) *Mucor miehei, M. pusillus*		3.4.23.6
Trypsin	Protese	Animal pancreas	None	3.4.21.4

[a]*Enzyme Nomenclature: Recommendations (1978) of the Nomenclature Committee of the International Union of Biochemistry*, Academic Press, New York. 1979.
[b]Usually a mixture of polygalacturonase, pectin methylesterase and pectate lyase.

ganisms are in wide use today, although it is now known that enzymes produced by these organisms are the actual agents responsible for the conversion of grapes to wine, milk to buttermilk or yogurt, etc. In addition to modern applications of ancient discoveries, enzymes extracted from living organisms also are widely employed in the food industry.

Enzymes used by food manufacturers are derived from edible and nontoxic plants, animals, and nonpathogenic, nontoxigenic microorganisms (*47*). Some of the enzymes used in food processing are given in Table 1 along with the sources of each. Because enzymes are catalysts, the amounts added to food (usually at an early or intermediate step in processing) represent only a minute fraction of the total food mass (*5*). Even this small amount may be reduced by further processing. For example, heating to produce desired organoleptic properties enhance shelf-life and ensure the absence of pathogenic microorganisms will denature or destroy the activity of most enzymes. The protein molecules which comprised the enzymes will still be present, but their physical shape will have been irreversibly altered by heating so that they no longer possess catalytic activity. There are also other methods of enzyme removal and/or inactivation such as raising or lowering the pH beyond limits which the enzyme can tolerate (*47*). Every enzyme exhibits a range of pH stability above or below which inactivation occurs. Many enzymes are inactivated by the acidity of the stomach.

The main organic constituents of foods are carbohydrates, proteins and lipids. It is often desirable to alter one or more of these constituents with enzymes during the conversion of raw to finished product. An important example of this involves the use of carbohydrases and isomerase to produce corn syrups from starch (*29, 32, 47*).

In one example of this conversion, alpha-amylase (IUB 3.2.1.1) first breaks long-chain starch molecules into shorter chains. Then glucoamylase (IUB 3.2.1.3) cleaves the individual glucose molecules from the chains. The resulting corn syrup has many commercial applications, but it is not as sweet as sucrose, the common table sugar obtained from sugar cane and sugar beets.

This deficiency of corn syrups has been overcome in recent years by the discovery of glucose isomerase (IUB 5.3.1.5), which converts glucose into fructose. The resulting high fructose corn syrup (HFCS) approaches the sweetness of sucrose and is less expensive. It is replacing the disaccharide in many applications.

There are many other novel and important applications of enzymes. For example, some foods and beverages do not store well in the presence of oxygen. By use of the enzyme glucose oxidase (IUB 1.1.3.4), which adds molecular oxygen to glucose to produce gluconic acid, it is possible to remove atmospheric oxygen safely and effectively from foods or beverages that are susceptible to oxygen.

Another interesting example is the production of juices from certain fruits and vegetables, where pectin content may become an important consideration (*47*). Pectin and pectic substances occur in plants. They are complex carbohydrates which are insoluble in water but nonetheless absorb water and, when dispersed, greatly increase viscosity. This is a desirable property for certain juices, such as those made from tomatoes, apricots and oranges, but the resulting lack of clarity is undesirable in apple and grape juices. Unfortunately, nature does not necessarily accommodate human taste. Raw apple and grape juice can contain considerable amounts of pectin even though most of us may not like them that way. For this reason, it is usually necessary to add pectic enzymes to raw apple and grape juices during processing to hydrolyze the pectin. Additionally, considerable amounts of juice can remain trapped in masses of pectic material. Through the use of pectic enzymes, such trapped juice can be freed. This makes juice extraction more efficient and economical, hence it lowers the price for consumers.

It is important to recognize that pectic enzymes (a mixture of three enzymes — see Table 1), as well as pectin, are naturally present in fruit juices, and where more enzyme activity is required, additional pectic enzymes may be added as indicated above. However, where high pectin content is preferred (e.g., apricot nectar, tomato and orange juices) the juice may be heated at an early stage in processing to denature native pectic enzymes and thereby

preserve natural pectin content. Another variation is used in jelly manufacture. Here, the native pectin is hydrolyzed by pectic enzymes, and then, after heating to denature the enzymes, commercial pectin possessing certain desirable properties is added to produce jelly of consistent quality.

Pharmaceutical/medical applications. Because of the great versatility of enzymes, their use is not restricted to food processing. Enzymes also have gained importance in the pharmaceutical/medical industry. For example, they are used in rapid and highly reliable clinical diagnostic tests. In one such test, the enzymes glucose oxidase and peroxidase (IUB 1.11.1.7) have been combined in a specific and sensitive assay for glucose in urine (a symptom of diabetes). The glucose oxidase/peroxidase test is superior to urine-glucose tests based on chemical reduction of glucose (*9, 25*). It has also recently been applied to the detection and quantitation of glucose in blood. Other enzymes which catalyze different reactions with glucose also are used in glucose determinations. Moreover, many physiologically important substances, such as blood urea nitrogen (BUN), triglycerides and glycerol, cholesterol, uric acid, and several physiologically important enzymes, can be rapidly and specifically assayed with commercially available enzyme-based tests.

Enzymes also are employed in antibiotic manufacture to alter the chemical structure of antibiotics and thereby increase the range of microorganisms which the antibiotics can control. A related and particularly interesting example is the therapeutic application of beta-lactamase (formerly penicillinase) (IUB 3.5.2.6), an enzyme which destroys penicillin. The gene which codes for penicillinase is found on certain plasmids (extrachromosomal DNA) and the acquisition of such plasmids by pathogenic bacteria confers penicillin resistance. However, the purified enzyme can also be used to treat people who are hypersensitive to penicillin but were inadvertently exposed to the drug (*47*). Thus, imaginative application has resulted in health benefit from an enzyme which functions in nature to the detriment of human health.

There are many other similar examples of the therapeutic uses of purified enzymes from pathogenic microorganisms, from the venom of poisonous snakes, from human urine and from a variety of other plant, animal and microbial sources (*19*). Enzymes may be used in the treatment of human maladies ranging from cancer and thrombosis to prevention of tooth decay (*19, 47*).

Enzyme detergents. The addition of enzymes to laundry products to aid in stain removal was developed by Rohm, who patented the idea in 1913. Various improvements were made on the original concept, and, by 1969, enzyme detergents claimed 50% of the market in Europe and almost 45% in the United States (*49*). Then, following widely circulated, unfavorable publicity concerning the possible development of allergies to enzymes inhaled as a result of dust formation, the use of enzymes in laundry products in the United States declined dramatically. However, an expert committee, with support from the United States Food and Drug Administration (FDA), has concluded that irritation from enzyme detergents does not exceed that of detergents which do not contain enzymes (*15*). In addition, methods have been developed to encapsulate enzymes in polymeric matrices which are too large to be dispersed in air as dust particles, yet retain enzyme catalytic activity in the laundry product. Hence, it is now possible to produce an essentially dust-free enzyme detergent (*49*).

The use of enzymes in laundry products offers prospects for decreasing energy (heating) costs as well as minimizing water pollution (diminishing the need for other chemical additives). Enzymes are being used widely and successfully in laundry products without evidence of adverse health effects in consumers (*49*).

Other uses. There are many other practical applications of enzymes. For example, enzymes are used widely in the textile and leather industries to remove undesirable substances from products during manufacture. Additionally, commercial enzyme preparations are available for use in septic tanks. Such preparations often contain many enzymes for decomposing complex carbohydrates, proteins and lipids, as well as viable microorganisms which use the enzyme-liberated products as nutrients and produce additional degradative enzymes to continue the cycle. Microorganisms producing appropriate enzymes are also used to detoxify pesticides, and other bacteria can remove nitrate and nitrite from water supplies (*47*). Certain microorganisms and their enzymes are gaining particular attention in the production of alcohol as fuel as well as in the production of food from inedible materials or by-products (*47*).

Future applications of enzymes

It is now apparent that additional useful and important applications of enzymes to societal improvement are limited only by the depth of our imagination and our resolve as a nation to encourage experimentation and innovation. Technological application of enzymology is a direct outgrowth of our scientific preeminence, and once reasonable safety has been established, new developments should be allowed to proceed unfettered. Many problems which disturb us and plague much of the rest of the world, such as unavailability of food, fuel, adequate medical and pharmaceutical supplies, clean water and pollution control, are amenable to enzyme technology. Enzymes are an immensely valuable renewable natural resource, and their imaginative use in improving human welfare should be nurtured.

By way of specific example, one area of great potential is enzymic nitrogen fixation. Nitrogen is an essential element for life [indeed, all enzymes contain about 16% nitrogen (*52*)], yet atmospheric nitrogen cannot be utilized by animals, plants and most microorganisms. Nitrogen can be "fixed" as ammonia (a biologically usable form of the element) by industrial processes which consume much energy (*31*). In contrast, blue-green algae and certain species of bacteria can produce ammonia from nitrogen and hydrogen in a much more efficient manner, although energy is still required (*52*). Hence, an important challenge is the

harnessing of the enzymic process of nitrogen fixation for industrial-scale production of ammonia. Such a development would go far towards alleviating global food shortages.

As the example given above illustrates, enzymes in the broadest sense are really inexpensive alternatives to energy-requiring physical processes, such as the application of heat and/or high pressure. This is because enzymes accelerate reactions which would proceed only very slowly, or not at all, under ordinary conditions. Moreover, because enzymes are so specific in the reactions which they catalyze, many important and highly useful chemical transformations could not be accomplished without them. For these reasons, the future of enzyme technology seems exceedingly important and bright.

MANUFACTURE, COMPOSITION AND CONSUMPTION OF ENZYME PREPARATIONS

Enzymes are manufactured because we need highly specific catalysts which are safe to use. Two considerations are of primary importance: (a) catalytic activity must be preserved during production and (b) the intended and proper use of enzyme preparations must pose no health risk for plant workers or consumers. These two central principles underlie enzyme manufacture and use.

Like all biological materials, enzymes are affected by the conditions under which they are produced and handled. Economically important enzymes are obtained from animals, plants and microorganisms. In the manufacture of enzymes there must be strict adherence to current Good Manufacturing Practices (CGMPs). (*8*).

Enzymes from animals

One of the first intentional developments by man of what could be called an "enzyme preparation" was rennet, a crude extract of the lining of the fourth stomach of ruminants. This extract contains various proteolytic enzymes which cause milk to curdle, a step essential for cheese production. Rennet is still obtained from this traditional source except that modern methods of enzyme manufacture and quality control are applied to ensure a product of consistent activity which is free of pathogenic bacteria and toxic substances (*3, 4, 8*).

Other crude enzyme mixtures are also obtained from animals at slaughter, such as pancreatin from the pancreas (contains several proteolytic, amylolytic and lipolytic enzymes), pepsin from hog stomachs, lipase from the throat glands of young ruminants and hyaluronidase from bovine seminal vesicles (used medically to facilitate the diffusion and adsorption of local anesthetics). An important perspective of enzyme production from animals is evident from the fact that in 1975, in the Federal Republic of Germany alone, pancreas glands from 13.3 million animals were required for the production of just 100 kg of pancreatin (*44*). As in the manufacture of calf rennet, high standards of quality are maintained throughout the production process to ensure the safety and efficacy of the final enzyme preparations.

Enzymes from plants

Enzymes of commercial importance are also obtained from edible nontoxic plants. The terms *edible* and *nontoxic* are both important, since some edible plants can contain toxic substances (e.g., potatoes and rhubarb) (*13*). However, the plants used for food enzyme manufacture are not known to produce or contain such toxins. Three plant proteases (bromelin, papain and ficin) are obtained, respectively, from the stalks of pineapple plants, the fruit of papaya and the sap of fig trees. Additionally, horseradish roots serve as the source of horseradish peroxidase (an important analytical and research enzyme), and barley seeds are the source of malt which contains amylase activity and is used in brewing (*47*).

Imported raw materials are surveyed for possible insect-derived contamination. If found, the product is processed to remove the contaminant. Another consideration common to all agricultural products is possible pesticide residues or mycotoxins in plant-derived enzyme preparations. Enzymes often are separated from other plant constituents by precipitation with organic solvents such as ethanol, acetone or isopropanol (*47*). Any organic toxins initially present are likely to be separated from the enzyme-containing protein fraction which precipitates.

Enzymes from microorganisms

Microorganisms are the most important source of commercial enzymes. Virtually any enzymic activity of industrial importance may be produced by one or more species of microorganism. This does not mean that microorganisms naturally synthesize animal or plant enzymes, but rather that microorganisms may produce their own enzymes to catalyze reactions that are also catalyzed by structurally different enzymes from animals or plants. Microorganisms are readily grown and manipulated on an industrial scale, and the synthesis of specific products, including enzymes by these organisms, can be regulated by using selected or genetically-engineered strains and/or varying growth conditions. Hence, the uniformity of composition of microbial enzyme preparations can be maintained.

Organism selection. Manufacturing a microbial enzyme begins with well-characterized pure cultures isolated from various sources. There are many cultures currently in use (Table 1). Microbial cultures used in food enzyme manufacture should have been tested to establish that they are nonpathogenic, nontoxigenic and do not produce antibiotics (*3, 4, 7, 45, 47*). Specific cultures often will have been subjected to many tests, and there should be little doubt that the microorganisms listed in Table 1, when handled under CGMPs, are safe for food enzyme manufacture. Cultures of the same or different species isolated anew from natural sources may also be of potential importance in food enzyme manufacture. The guidelines and procedures which we present below can be used to assess the safety of new isolates.

A culture (currently in use or isolated anew) will have been selected on the basis of its ability to synthesize a desired enzyme. However, the enzyme may be produced at

only relatively low rates. Moreover, the culture may also produce other undesired enzymes. For example, microbial rennet preparations often contain unwanted enzymes which can produce off-flavors in cheese on prolonged aging (*47*). Hence, it is common practice to attempt to improve the desirable qualities of the isolate by altering growth conditions, usually in conjunction with strain selection by mutation or other types of genetic manipulation. The result can be a special strain that will not survive in nature but is very useful from a commercial standpoint.

Laboratory-generated mutant strains characteristically lack certain functional or regulatory properties.[2] While the primary structures of proteins can be altered within limited ranges by mutagenesis, mutants possessing enzymes with improved catalytic activity for their normal substrates have not been reported (*30*). Moreover, no one has ever reported a mutation which transformed an otherwise nontoxic enzyme or protein into a toxin. It is now possible to introduce foreign genes into microorganisms by using DNA cloning techniques so that entirely new proteins are produced, but this should not be confused with mutagenesis where the intrinsic DNA of an organism is altered.

A useful mutant strain might be one which has lost a regulatory function that limits the synthesis of a desirable enzyme so that the mutant cannot stop synthesizing the enzyme and continues to produce it in great excess of biological need. The mutant may also have lost the ability to synthesize one or more unwanted enzymes. Additionally, it may have been manipulated genetically so that more than one copy of the gene coding for the desired enzyme is present, hence, there are more "blue-prints" available (*47*). Such organisms are really genetically impaired and are maintained in the laboratory or industrial setting by using specific, well-controlled growth conditions. These microorganisms have not been found in nature probably because they cannot compete successfully with the wild-type (nonmutant) parent or other microorganisms. It is also important to note that when the parental isolates are pathogenic, the derived mutant strains are characteristically less hazardous. Of course cultures used for food enzyme manufacture are not pathogenic, but by way of example, mutant strains of *Salmonella typhimurium* developed for routine mutagenesis testing are far less virulent than *S. typhirmurium* found in nature (*1*). Therefore, in choosing innocuous isolates for enzyme production, the process of enzyme manufacture from microorganisms becomes inherently safer.

The nonpathogenic, nontoxigenic microbial cultures traditionally used in enzyme manufacture are also ideal candidates for cloned DNA. For example, the gene for a useful enzyme that is not synthesized by *Bacillus subtilis* could be introduced into the organism. The new "strain" would then produce the new enzyme product and would not present a pathogenic or toxigenic risk greater than that of its "parents," the nonpathogenic *B. subtilis* and the gene for the useful enzyme.

Large-scale growth. There are two ways to grow microorganisms on an industrial scale. One way is to use liquid medium which is agitated and aerated, and the other way is to use solid or semi-solid medium held in large trays or drums (*16, 47*). In both cases, it is necessary to control environmental factors such as temperature, pH and degree of aeration. Equipment must be designed for easy cleaning and sterilization. Conditions must be employed which minimize the growth of contaminating microorganisms that will ruin the fermentation. During growth, cultures are routinely sampled and tested for possible contamination (*16, 47*).

All ingredients used to formulate the growth medium should be free of toxic contaminants (*7, 8, 16, 45, 47*). It is important that any "carry-over" of growth medium into the final enzyme preparation not bring with it possible toxic substances, especially when the enzyme being manufactured is intended for food processing.

Enzyme extraction, concentration and standardization. The desired enzyme may be present in the medium or inside the cells. Enzymes secreted into solid or semi-solid medium, and most intracellular enzymes, are extracted before further processing. In this context, extraction means to "wash out" and solubilize the enzyme in an aqueous solution (*16, 47*). Where the enzyme is secreted into a liquid growth medium, an extraction step is not necessary.

Enzymes secreted into solid or semi-solid media may be extracted directly into water solutions using a counter current system which filters as well as extracts (*16, 47*). Alternatively, solid or semi-solid media containing the microorganisms may be dried, ground and treated with water solutions to solubilize the desired enzyme. This method can be used to recover both intra- and extracellular enzymes. In the case of intracellular enzymes from microorganisms grown in liquid media, the cells are first collected by centrifugation or filtration and then ruptured by any of a number of physical and/or chemical procedures (*16, 47*). The enzymes are then extracted from ruptured cells with aqueous solutions.

After extraction, enzyzme solutions are usually concentrated to reduce volume. It is common to use ultrafiltration to reduce the amount of water and substances below specified molecular weights (e.g., salts, small organic molecules and peptides). Sometimes enzymes are concentrated by precipitation with salts or organic solvents, but because of organic solvent cost this method is not as common today as it was 10 years ago (*47*). In other cases, con-

[2]Under certain conditions an inducible enzyme can be made constitutive by mutation in the regulator, operator or (more rarely) the promoter region of the genetic operon. The enzyme will then be expressed in the absence of the inducer. Thus, under fermentation conditions used to produce an enzyme, production of "new" enzymes or proteins can be made to occur. These proteins or enzymes were originally present in the genetic material of the parent and would be normally synthesized under the right fermentation conditions without mutation. In addition, mutation induces minor changes in base sequence of DNA encoding for proteins and enzymes (base change, deletion, etc.). Thus, minor changes in protein structure are possible as a result of mutations affecting the structural gene. These changes can lead to increased enzymic activity or they may decrease or destroy enzymic activity (*18*).

centration is accomplished simply by removing water through evaporation. Preservatives are almost always added during processing, and optionally in the final preparation, to prevent microbial growth and to stabilize and maintain the desired enzymic activity. Proper and appropriate use of preservatives and stabilizers serve to protect the consumer from unsafe or ineffective enzyme products (*7, 8, 16, 47*). When the enzyme is intended for addition to food, all such additives and diluents must be acceptable to the FDA for use in food. They must be of food grade quality and the levels used must not exceed specified limits.

Most industrial enzymes are not purified to any significant extent because purification is not necessary to achieve safe and useful products (*3, 4, 16, 47*). However, it is sometimes desirable to remove or destroy unwanted enzyme activities which would otherwise interfere with effective use of the desired enzyme preparation. For example, rennet produced by some microorganisms contains lipase activity which will make the finished cheese rancid. By carefully exposing the crude rennet to heat or low pH, the lipase can be inactivated without affecting the protease activity. In this example, the unwanted lipase is not physically removed (as in purification); the protein remains but is no longer catalytically active (*47*). Because of expense, physical separation normally is accomplished only when there is a market for the individual separated enzymes, although some manufacturers do highly purify certain enzymes of particular economic importance. For example, one company produces a very pure, crystalline glucose isomerase preparation for its own use (*47*).

Following extraction, concentration and stabilization, enzyme preparations are standardized (*3,4,47*). Because enzymes are catalysts, they are marketed in terms of units of catalytic activity rather than by weight or volume. A unit of catalytic activity for an enzyme preparation is defined in terms of the transformation of a given amount of substrate during a specified period of time under stated reaction conditions. Biochemists often use a unit defined by international convention, which is the amount of enzyme required to transform one micromole of substrate per minute under specific reaction conditions. However, this definition is not applicable to many commercial uses where the substrate is part of food (e.g., Swift's hamburger test for papain; *47*). Hence, many assays for industrial enzymes are based on specific application rather than uniform convention.

The standarization procedure consists of using a specific quantitative assay to determine the level of enzyme activity per milliliter or gram of the final enzyme preparation and then adjusting the activity (usually by dilution of the enzyme preparation) to conform with a desired level of activity which is convenient to use. Unstandardized enzyme preparations may also be sold, and, in this case, total activity is stated and will vary between lots.

Given that enzymes are marketed on the basis of activity rather than weight or volume per se, it follows that the activities and amounts of other enzymes, as well as the levels of nonenzymic catalytically inert materials, may vary from lot to lot and almost certainly from source to source (*47*). Moreover, since enzyme preparations are almost always relatively crude mixtures, it is apparent that anything produced by the source organims, and anything purposely or inadvertently introduced into the system during enzyme manufacture, may end up in the final enzyme preparation. For this reason, it is important that the source organism not produce or contain toxins. To avoid inadvertent contamination with unsafe substances, it is necessary that CGMPs be followed during enzyme manufacture. There are strict limits on the levels of heavy metals which will be tolerated, and there are requirements for demonstrating microbiological safety (absence of salmonellae, etc.) (*3, 4, 16, 45, 47*).

Immobilized enzymes

Some enzymes are sold in an immobilized form, i.e., products containing enzymes that have been immobilized by adsorbtion, entrapment, reaction with cross-linking agents or covalent attachment to insoluble supports (*29*). The safety evaluation of products such as these may require consideration of factors other than the safety of the enzyme, its source and the by-products of the production methods. For this reason, safety evaluation of immobilized enzymes will not be included in this paper.

Consumption levels

Total Organic Solids (TOS). Enzymes are marketed by units of activity rather than by weight or volume, and enzyme preparations always contain other substances (salts, preservatives, stabilizers, carriers, nonenzymic organic material, etc.) (*16, 45, 47*). Further, some enzymes are added to food and remain there, although they may be inactivated by heat or other treatment in the finished food product. On the other hand, some enzymes only come in contact with the food (immobilized enzymes) but do not stay there. For these reasons, it is not an easy matter to estimate total enzyme use and consumption.

The most logical means currently available for arriving at a reliable estimate of enzyme use and consumption was developed by the Ad Hoc Enzyme Technical Committtee (AHETC), a trade group representing companies that produce or distribute enzymes for food use. AHETC set forth the concept of Total Organic Solids (TOS; *5*) as a means of determining the toxicological significance of material derived from the enzyme source. TOS is defined as the sum of the organic compounds, excluding diluents, contained in the final enzyme preparation. It is derived experimentally as follows:

$$\text{TOS}\,(\%) = 100 - \text{A} - \text{W} - \text{D}$$

where A = % ash contained in the extract or isolated enzyme, W = % water in the extract or isolated enzyme, D = % diluents (if any, or carrier if enzyme is immobilized).

The 1978 Enzyme Survey. The Food and Nutrition Board (FNB) of the National Research Council's Assembly of

TABLE 2. *Selected enzymes and their maximum use in various foods based on TOS (Total Organic Solids) (5).*

Enzyme	Food category	Maximum use[a]
Papain	Baked goods	0.0078%
	Meats/meat products	0.0044%
	Beer/ale/malt beverages	0.0045%
Rennet (and other milk clotting enzymes)	Cheese	0.036%
	Gelatins/puddings/custards	0.0040%
Bromelain	Candy	0.000016%
	Fats and oils	0.000084%
	Snack foods	0.00056%
Pectinase	Baked goods	0.00000026%
	Fruits/juices	0.0035%
	Non-creamed soups	0.060%
Invertase	Candy	0.0078%
α-Amylase	Breakfast cereals	0.0030%
	Sugars/frostings	0.052%
	Gelatins/puddings/custards	0.0000020%
	Corn syrup	0.052%

[a]Percent of food based on TOS.

Life Sciences has undertaken several surveys of industrial use of food additives. In 1977, the FNB's Committee on GRAS List Survey — Phase III was asked by the FDA to organize an extensive survey of enzyme use in food processing. The Committee worked closely with AHETC and the FDA in developing questionnaires; then the AHETC distributed the survey forms to users and manufacturers of enzymes on a confidential basis. The FNB Committee received the completed forms directly for the respondents, reviewed and analyzed the data, and submitted a report to the FDA. The document is entitled *The 1978 Enzyme Survey (5).*

The survey report contains extensive information on 23 enzymes and an analysis of their use in a detailed list of specific food items. Average and maximum use levels are estimated by TOS. Removal and inactivation of the enzymes by further processing is also tabulated. Table 2 contains some examples from this survey demonstrating the low levels at which enzymes are added to foods.

ENZYME SAFETY

Current status

Exhaustive literature reviews commissioned by the FDA for food enzymes from microbial (*43*) and nonmicrobial (*11, 44*) sources support the proposition that enzyme preparations from nontoxigenic, nonpathogenic organisms are safe to consume. This conclusion is strengthened by the report of the Joint FAO/WHO Expert Committee on Food Additives, which evaluated both published and unpublished data (*12*). There are numerous GRAS affirmation petitions currently before the FDA which also contain safety data on enzyme preparations (*46*).

It is not surprising that the enzymes used in food processing have proven to be nontoxic when tested in animals. In fact, very few toxic agents have enzymatic properties and those that do, e.g., diphtheria toxin and certain enzymes in the venoms of poisonous snakes catalyze unusual reactions which are completely unrelated to the kinds of catalytic transformations that are desirable in foods. Hence, the only relevant issue is whether enzyme preparations contain toxic contaminants. It follows that, if the source organisms do not produce toxins and if CGMPs are followed during manufacture, then the resulting enzyme preparations will not contain hazardous materials.

In practice, industrial enzymes have a strong record of safe use in food processing. However, as with all food components, it is important that scientific data be provided to show that enzyme preparations, particularly those lacking a long history of safe use, are safe to consume. To develop a logical approach to this issue, we shall first consider the factors which bear on the safety of enzymes and then present guidelines for assessing enzyme safety.

Safety considerations

Safety of source organism. The safety of the source organism should be the prime consideration in assessing the probable degree of safety of an enzyme preparation intended for use in food. For example, if the source organism is a food animal, an edible and nontoxic plant, or a nontoxigenic and nonpathogenic microorganism which does not produce antibiotics, then it follows that enzyme preparations obtained from that source organism using CGMPs (*8*) will be safe to consume at the low levels encountered in processed foods. Moreover, in other instances

where toxic contaminants are present, they may be removed during manufacture.

With regard to microorganisms used in enzyme manufacture, we have discussed previously our contention that mutagenesis in the laboratory does not result in the acquisition of new genes, so it is not possible for an isolate to acquire a new toxin gene by mutation. It may be theoretically possible for a mutation to alter the structure of an otherwise nontoxic enzyme in such a way that the enzyme becomes toxic (*10*), but there is no experimental basis for this notion and we consider it to be remote. Advances in DNA sequencing may ultimately be useful in providing definitive proof of nontoxicity.

Proving that a new microbial isolate does not produce a toxin elaborated by other strains in the same species is complicated by the fact that toxin production may be affected by growth conditions. Under some conditions, toxin synthesis may be high, whereas under other conditions, it may be low or undetectable. Hence, to establish that an isolate is nontoxigenic in an absolute sense may not be possible strictly from data on toxin expression. By assaying toxin production under a variety of growth conditions, the probability of demonstrating toxigenic potential is increased. Moreover, if an isolate is grown under conditions where other closely related organisms elaborate a toxin, the reliability of a negative result is strengthened even further.

In practice, enzyme preparations will not contain all of the substances that a source organisms is able to produce. For example, enzymes which are concentrated by ultrafiltration or precipitation will contain far fewer low molecular weight components than are present in crude enzyme extracts. For this reason, even if an organism produces low levels of a potentially hazardous substance, the amount of a finished enzyme preparation needed to produce a deleterious effect in animals likely will be far above the low concentrations at which enzyme preparations are employed in food processing. Published animal feeding studies and summaries of unpublished experiments reviewed by expert committees (*12, 43, 44*) fully support this conclusion.

Pathogenicity. If an isolate is known to be or suspected of being a human pathogen, it will almost certainly not be further considered for commercial enzyme production unless it is the singular source of a unique and useful enzyme. The problems inherent in maintaining and handling cultures of pathogenic organisms on an industrial scale make it unlikely that they will ever be used in the manufacture of enzymes for food processing, and there are federal regulations concerning this issue (*7*). However, high purified enzymes from pathogenic bacteria are produced commercially and used with medical supervision in the treatment of disease (*19*).

Carcinogens and mutagens. No one has ever reported an enzyme which when fed was mutagenic or intitated carcinogenesis.[3,4] Given our current understanding of the processes of carcinogenesis and mutagenesis (*34, 51*), it is implausible to expect that the protein component of an enzyme or protein with such activity will ever be discovered[4]. Rather, attention should be directed towards the relatively small organic molecules (in general, MW <500 daltons) that possess carcinogenic or mutagenic activity and which might reasonably be expected to contaminate a given enzyme preparation.

Enzymes from mammals commonly used as food in the United States will not contain mutagens or substances which can initiate [3] carcinogenesis as long as CGMPs are followed. Some plants are known to produce carcinogens (*13, 34*), but the pineapple, fig, barley and papaya are not among them. The fungal and bacterial enzyme sources listed in Table 1 also are not known to produce carcinogens or mutagens. However, fermentative yeasts, such as *Saccharomyces cerevisiae*, may produce low levels of urethan (*37*), a carcinogen which is not mutagenic in the Ames test (*1*), as a natural by-product of fermentation. For this reason bread, wine and beer often contain low levels of urethan (*37*). There are no reports of urethan in yeast enzyme preparations. Moreover, where yeast enzyme preparations are concentrated by ultrafiltration or precipitation, small molecular weight compounds, such as urethan, will be removed or greatly decreased in concentration. For this reason it is unlikely that urethan levels in yeast enzyme preparations would exceed the levels found naturally in bread, wine and beer.

Several long-term animal studies (>90 days) have been conducted with enzyme preparations from microorganisms, and none showed evidence of carcinogenicity or chronic toxicity (*12, 43*). It is necessary to conduct such long-term tests for each new microbial culture, or for each new enzyme? We think not. For example, we have been unable to locate a single confirmed report of a carcinogen or mutagen produced by bacteria, other than certain *Actinomycetales*, particularly *Streptomyces*, when grown in ordinary culture media. When nitrite and secondary amines are added to culture media, a few bacterial species appear capable of generating nitrosamines through unknown mechanisms (*35*). However, there is no reason for nitrite and secondary amines to be added to culture media intended for use in food enzyme manufacture. Nitrosamines,

[3]It is important to recognize that the process of carcinogenesis as now understood consists of two stages. The first stage is called *initiation*, the second *promotion (39)*. Some animal products, e.g., certain fats and hormones, may at high doses and in certain well-defined experimental systems promote specific types of cancers. However, it has not been shown that these substances can initiate cancer, and it is commonly accepted among experts in this field that they are not complete carcinogens. Animals exposed to carcinogens may metabolize them to other forms which retain carcinogenic activity, e.g., aflatoxin M_1 in the milk from cows exposed to aflatoxin B_1 in their diets; (*42*). Animals may also generate nitrosamines from nitrite and secondary amines in their gastrointestinal tracts (*35*). However, mammals are not known to produce substances as normal body constituents which experts would classify as carcinogens.

[4]It is possible for certain enzymes that act on nucleic acids, such as DNA-dependent DNA ploymerase, to be altered by mutation in such a way as to become error-prone, thus resulting in further mutation in the organism containing the error-prone polymerase (*48*). However, such enzymes would not be produced for use in food processing. Moreover, should such enzymes be present in food enzyme preparations, they would almost certainly not enter human cells and produce an adverse effect. They are also produced by some *Streptomyces* sp. antibiotic proteins with mutagenic and DNA-damaging activities due to the presence of nonprotein prosthetic chromophores, i.e., the apoproteins themselves are without such activity (*25a, 39a*).

or any other classes of carcinogenic or mutagenic chemicals, should not be considered either a real or potential problem area in enzyme manufacture from bacteria (other than certain *Actinomycetales*).

In contrast, some antitumor agents and antibiotics produced by *Actinomycetales*, particularly certain *Streptomyces*, are weakly carcinogenic, e.g., azaserine (*34*). Moreover, some mycotoxins have carcinogenic and mutagenic activities (*33, 34, 42*). If there is reason to believe that such substances might be produced by a new culture under test, then specific chemical, biochemical or biological tests for the substances should be conducted.

Teratogens and reproductive effects. Various dietary deficiencies and excesses, hormones, drugs, agricultural and industrial chemicals, naturally-occurring toxins, and physical and biological agents produce, under some circumstances, teratogenic effects or reproductive deficiencies in experimental animals (*20, 27*). Some of these agents or conditions, such as German measles, alcohol abuse, and certain drugs and antibiotics, produce similar effects in humans. However, enzymes are not among the substances which have been shown to cause teratogenesis or reproductive deficiency. In fact, in a four-generation study in rats, a rennet preparation from *Mucor pusillus* produced no evidence of teratogenicity or toxicity towards the reproductive system (*12*), and similar negative data have been obtained for various enzymes from other microbial (*43*) and nonmicrobial (*11*) sources. Those microbial metabolites which could pose such a risk should be detected either as certain specific antibiotics (*20, 27*) or as acute/subchronic toxins (*42*).

Antibiotics. Antibiotics are chemicals produced by various species of microorganisms which kill or inhibit the growth of other microorganisms. They are really a special class of toxic agents which are useful to man in the control of disease. It is well-documented that a sensitive microorganism can acquire plasmids which confer antibiotic resistance on the host (*40*). For this and other reasons enzyme preparations intended for use in food processing should not contain antibiotics. There are methods for assessing enzyme preparations for antibiotic activity (*4*).

Allergies and primary irritations. Industrial enzymes are foreign (nonhuman) proteins, and as such, may be allergenic for humans under certain conditions. The group most likely to be affected are plant workers (*11, 15, 47, 49*). There are methods and procedures for protecting workers from this potential hazard and it is considered to be a manageable problem (*15, 47, 49*).

There are no confirmed cases of allergies or primary irritations in consumers caused by enzymes used in food processing. This is probably due, in part, to the low levels of enzymes added to foods. Foods naturally contain a wide variety of foreign (nonhuman) proteins, many of which are present at levels far higher than the industrial enzymes added as processing aids. Allergies and primary irritations from enzymes used in food processing should be considered a low priority item of concern except in very unusual circumstances. There is no justification for requiring routine testing of enzyme preparations for allergic responses or primary irritations relative to consumer safety.

Toxins involved in food posioning. A few bacterial species produce toxic proteins or peptides which can cause food poisoning. These include both enterotoxins and neurotoxins (*41*). There are immunological assays or animal systems for detecting such toxins. Within a bacterial species known to cause food poisoning via a toxin, usually only some, but not all, strains produce the toxin. Hence, nontoxigenic strains can be isolated (*41*). Some bacterial toxins are actually coded for in bacteriophage DNA which has become integrated into the bacterial genome as a prophage. "Curing" the organims of the prophage results in loss of toxicity (*41*).

Bacterial toxins which cause food poisoning are, by definition, substances which produce acute toxic responses following introduction into the gastrointestinal tracts of sensitive annimals. The nature and severity of the toxic response may vary among animal species under test, as well as the amount of toxin required to produce a measurable effect.

Products of enzymic reactions. Enzymes are used in food processing because they produce desirable changes in the natural food constituents. They are usually inactivated or removed before the final food product is marketed. As such, enzymes should be classified as *processing aids* or *secondary direct additives*. Declaring their presence on the label of a food product, in most cases, would be incorrect, since only rarely is the active enzyme present in the final product. This unique status of enzymes can lead to a new question, however. Are the products of the enzymic reaction safe? Developing an answer to this question requires an understanding of what the enzyme is doing in producing an apparently favorable transformation in the food.

Most of the enzymes used in food processing are degradative enzymes which split macromolecules, i.e., proteins, complex carbohydrates and lipids, into smaller subunits. Another important example is glucose isomerase, which catalyzes the conversion of glucose into its isomer fructose. Both glucose and fructose are nutritive and nontoxic. Only one enzymic reaction used in food processing is known to yield a potentially toxic product. Pectic enzymes increase the methanol content of treated fruit products, but the amount produced is far below the hazard level (*47*). There are reliable and rapid assays for methanol in food.

The question of hypothetical, potentially hazardous enzyme reaction products is difficult to evaluate, but probably its importance is marginal. For example, proteases from all sources degrade proteins into peptide fragments and amino acids. However, different proteases attack proteins at different sites and may produce different sets of peptide fragments from the same protein substrate (*52*). There are many biologically active peptides in nature which serve in various metabolic regulatory capacities. One may wonder if the peptides produced by proteases have any biological properties of their own. Until recently, most biochemists would have considered as highly remote the possibility that toxic peptides might be generated from

otherwise nontoxic proteins, and, indeed, it should still be considered speculative. However, a recent report (*53*) indicates that peptides with neuropharmacological properties are generated by the action of the natural animal digestive enzyme, pepsin, on wheat gluten or casein, i.e., the major protein of milk. The peptides are called "exorphins" because they mimic in vitro the action of opioid-like peptides, the endorphins, which are produced naturally by animals. It is suggested that such peptides may form during digestion of some food proteins in the human gastrointestinal tract, and could have physiological significance (*53*). The possibility of such peptides forming in processed foods treated with proteases was not considered.

This example illustrates the difficulty that arises when one attempts to establish absolute safety. Such a goal would be extremely difficult for a static system, and is clearly impossible when dynamic forces, such as basic scientific inquiry, continually expand our understanding and knowledge. However, there is also no reason, on the basis of available information, to fear that processed foods treated with proteases might pose a hazard, especially one that is greater than that posed by our own digestive systems. This is clearly a research area which deserves further support, especially as it relates to human physiological significance and development of specific and relevant assays.

Interactions between enzymes and other food components. It is well-known that certain drugs are not compatible with one another and that combinations of such incompatible drugs can result in interactions which are toxic (*28*). It has been suggested that such interactions might also occur between enzymes and other components of beverages or food products (*6*). However, there is no scientific basis for such speculation. It is extremely unlikely that enzymes, which are used at very low concentrations and are almost always inactivated or removed before the finished food or beverage is marketed, could produce a toxic effect due to interaction with another substance. Given the high specificity of enzyme action, it is difficult to imagine such an occurrence. The highly improbable possibility of toxic interactions involving food enzymes should not be afforded serious consideration unless supporting data appear in respected and well-refereed scientific journals.

Direct effects of food enzymes on consumers. Under the usual conditions of use in foods, enzymes do not pose a hazard for consumers. For example, ingesting an active protease at relatively low levels could hardly affect the human gastrointestinal tract, where many potent proteases, such as trypsin and pepsin, already are present at levels sufficient to digest food. This view is supported by the report of an expert committee (*11*). Proteases may adversely affect the skin, mucous membranes of the nose and throat, and lungs, and such effects are sometimes seen workers who handle large quantities of proteases. However, such occurrences are extremely rare in consumers who use much lower levels of active enzyme (*11, 15*), and it is not possible for heated foods containing inactive proteolytic enzymes to pose such a threat. Active proteases are, of course, widely distributed in fresh fruits, vegetables, cheeses and other uncooked foods which may be consumed.

We know of no reported adverse effects on humans from lipase/esterases or carbohydrases in foods. Moreover, many enzymes are inactivated in the gastrointestinal tract and digested as protein.

Concept of relative safety

The terms *nontoxigenic* and *nonpathogenic* should not be considered in an absolute sense. In the real world they are relative concepts which convey certain probabilities. A nontoxigenic organism is one which does not produce injurious substances at levels that are detectable or demonstrably harmful under ordinary conditions of use or exposure. In the same vein, a nonpathogenic organism is one that is very unlikely to produce disease under ordinary circumstances. Thus, *Aspergillus oryzae* should be considered nontoxigenic because it does not produce detectable levels of aflatoxin (*23, 50*) and is not listed with molds known to produce other mycotoxins (*42*). Strains in commercial use did not produce detectable levels of beta-nitropropionic acid (*36*) and there are no reports of this organism producing adverse effects in animals. Likewise, *S. cerevisiae* should be considered nontoxigenic even though low levels of the carcinogen urethan are produced during fermentation (*37*) because, as far as we can tell, the amount of urethan is too low to be significant. Applying an absolute definition in this case would result in the banning of bread, wine and beer. There is no reason to believe that such an extreme measure would make our lives safer! As long as the levels of urethan in fermentative yeast enzyme preparations do not exceed those found in fermented foods and beverages, they should not be a cause of concern.

Aspergillus niger produces low levels of toxic substances (*22*), but it is only after such substances are extracted and concentrated that toxicity can be demonstrated. This example points up the important distinction between *toxin*, a chemical entity, and *toxic effect*, a biological phenomenon produced by toxins only at effective doses. Synthesizing low levels of toxins per se should not be sufficient to support redefining *A. niger* as a toxigenic organism, and it should remain classified as nontoxigenic. In the same way *B. subtilis* should be considered nonpathogenic even though one could imagine an individual with an extremely compromised immunological system succombing to a *B. subtilis* infection. Under more ordinary circumstances, *B. subtilis* does not cause disease.

These concepts are important in considering safety assessment. Absolute safety is not achievable and cannot be our goal. Rather, we should think in terms of probabilities tempered with common sense.

Animal testing for toxins

The purpose of animal testing is to assure that toxic effects are not produced by non-enzyme substances in enzyme preparations under realistic projections of use. There

is no basis for concern that the enzymes under consideration in this report are themselves toxic. Acute and subchronic oral toxicity studies (to be proposed) should be conducted with two animal species (*24*). This is necessary to compensate for possible species variation in toxic response. For example, rats are much more sensitive to aflatoxin B_1 than mice, whereas dogs are more sensitive than rats to ochratoxin A (*42*). There are also species variations in response to the protein/peptide enterotoxins and neurotoxins of bacteria (*41*). Additionally, some animal species are capable of emesis, e.g., dogs and pigs, whereas others are not, e.g., rodents. Selection of appropriate test animals should be based on two criteria: (a) which toxins could be produced by the source organism and (b) which toxins have already been eliminated from further consideration by the use of specific chemical/biochemical assays. In many instances, rats and dogs may be the most appropriate test animals (*24*).

Guidelines for determining enzyme safety

Basic premises. In developing guidelines to assure the safety of enzymes used in foods, we have adopted the following basic premises to guide our thinking. The rationale for each of these premises can be found in preceding sections.

1. Enzymes are naturally occurring proteins. Only a very few, highly unusual enzymes are toxic and they would not be used in foods.
2. There is no basis for concern that enzymes acting on otherwise wholesome food constituents will generate harmful products. Hence, there is no reason to test enzyme-treated foods for toxicity.
3. New enzymes could be derived from animals, plants or microorganisms. However, for technical reasons it is likely that most new enzyme preparations will be derived from microbial sources, in many instances new microbial species or strains.
4. Enzymes are added to food at very low levels. Failure to demonstrate harmful materials in, or toxic effects from, concentrated enzyme fractions, which when diluted yield finished enzyme preparations for marketing, gives reasonable assurance of their safety. Alternatively, failure to demonstrate harmful materials in, or toxic effects from, cultures or crude extracts of a proposed source microorganism, gives reasonable assurance of safety for any enzyme preparation which may be produced from that source organism using CGMPs.
5. If a microbial culture does not produce known toxins and if its metabolites are nontoxigenic in the sense that they do not produce food poisoning, intoxication or illness when ingested, then enzymes derived from that culture using CGMPs will be safe for use in food processing.
6. If there are toxigenic strains of the species to which the new culture belongs, then growth conditions under which those strains produce toxins should be tested. The condition(s) to be used for enzyme manufacture would, of course, be included. It is also prudent to test mutants for toxins produced by other *strains* of the same *species* even if the parent culture is negative for such substances.
7. Certain microbial species produce antibiotics, which are detectable in appropriate bioassays.
8. Some of the filamentous fungi and *Actinomycetales* produce toxins. A few of these substances are carcinogenic, e.g., aflatoxin, and some also possess antitumor and antimicrobial activity, e.g., azaserine. Such metabolites may be detected with specific chemical, biochemical or biological assays.
9. Bacteria other than *Actinomycetales* may also produce acute toxins. Of specific concern are the peptide/protein toxins that act via the oral route, e.g., enterotoxins and certain neurotoxins. Toxins associated with foodborne illness can be detected with serological or animal assays.
10. Bacteria as a group (other than *Actinomycetales*)are not known to produce carcinogens or mutagens when grown in ordinary culture medium which does not contain nitrite and secondary amines.
11. Yeasts as a group are not known to produce toxins, although some yeasts are pathogenic. The carcinogen urethan may form at very low levels in yeast fermentations. Urethan can be detected by chemical assay.

Microbial enzymes. Guidelines for determining safety of microbial enzymes are shown in Table 3. These guidelines may be applied to concentrated enzyme fractions which are diluted to produce finished enzyme preparations. Alternatively, the guidelines may be applied to crude culture extracts or whole cultures from which enzymes are manufactured. If the crude culture extracts or whole cultures are judged to be safe, then enzymes can be manufactured from these sources without further testing.

It is important to note the following features concerning the guidelines in Table 3.

1. All test materials must be evaluated for antibiotic activity.
2. No test material can pass through the Decision Tree without being tested for toxic constituents.
3. Two animal bioassay systems are proposed. The first is a single oral challenge. The purpose of this assay is to evaluate the test material for food poisoning toxins, specifically enterotoxins and certain neurotoxins, which are protein or peptide toxins produced by a few bacterial species. The second proposed bioassay is a subchronic feeding study in two appropriate animal species. The purpose of this procedure is to detect mycotoxins and other toxic substances which might not produce acute toxicity. All

TABLE 3. *Guidelines for determining the safety of microbial enzymes*[a].

A. *Decision Tree*	If yes --proceed to--	If no
1. Is the test material free of antibiotics?[b]	A,2	D
2. a. *For bacteria and yeast,* is the test material:		
i. Free of toxins[c] known to be produced by other strains of the same species?	A,3	D
ii. If there are no known toxins[c,d] produced by other strains of the same species, is the no-adverse effect level in a single oral challenge at least 100 times greater than the estimated mean human consumption level?[e,f]	B	D
b. *For molds,* is the test material free of detectable levels of aflatoxin B_1, ochratoxin A, sterigmatocystin, T-2 toxin, zearalenone and any other toxins known to be produced by strains of the same species?[g]	C	D
3. Is the no-adverse effect level in subchronic (90-d) feeding sutdies at least 100 times greater than the estimated mean human consumption level?[f,h]	ACCEPT	D
B. *Special considerations for certain yeasts and bacteria*		
1. If the source culture is well-known, widely distributed, nonpathogenic yeast, e.g., certain species of the genus *Saccharomyces,* or if it belongs to a bacterial species that is well-characterized, commonly present in foods, has a history of safe use in food enzyme manufacture, and has never been implicated in foodborne disease, e.g., *Bacillus coagulans, Bacillis licheniformis, Micrococus lysodeikticus,* and *Bacillus subtilis (17),* the test material can be ACCEPTED at this point.		
2. Test material from other bacteria and yeasts must be considered under part A,3.		
C. *Special considerations for certain molds*		
1. If the source culture is well characterized, commonly present in food, has a history of safe use in food enzyme manufacture, and has never been implicated in foodborne intoxication or disease, e.g., *Aspergillus oryzae, Apergillus niger* and *Rhizopus oryzae (16,23,36,41,42,43,45,47,50),* the test material can be ACCEPTED at this point.		
2. Test material from all other species of molds must be considered under A,3.		
D. *Disposition of materials that fail any Decision Tree requirement*		
A negative answer to questions 1, 2 or 3 signifies the presence of an undesirable substance and the material is not acceptable for use in food. If the undesirable substance can be removed, the purified material must be passed through the system again beginning at the point of the original negative answer.		

[a]These guidelines are intended for crude culture extracts, for whole cultures, and for concentrated enzyme fractions which, when diluted, become enzyme preparations suitable for marketing.

[b]As determined by (*4*) or comparable methods.

[c]For the purposes of these guidelines, the term "toxin" refers to a substance which is regarded by experts as a cause of food poisoning, intoxication or illness when ingested. Examples are staphylococcal enterotoxins, botulinal neurotoxins and mycotoxins.

[d]Certain cultures in this category are acceptable on the basis of a single acute oral toxicity test, as explained in part B,1. Cultures that fall under part B,2 can go directly to part A,3 without an acute oral toxicity test. This is permissible because the subchronic feeding specified in part A,3 is more rigorous and more meaningful than the acute oral toxicity test embodied in part *A,2,aii*.

[e]Expressed as mg/kg body weight and determined using two appropriate animal species.

[f]Estimated mean consumption level is calculated from the sum of the intakes for each food category in which the material is expected to be used. An example of such determination is: (USDA mean portion size) × (Maket Research Corporation of America eating frequency for the entire population) × (the usual level of use expressed as TOS for the enzyme in question)(*2,14*).

[g]As determined by (*38*) or comparable methods.

[h]Expressed as mg/kg body weight/day, and determined using two appropriate animal species.

known microbial toxins active via the oral route and present at effective levels will be detectable by these procedures. It should be pointed out that preparations will be tested in these proposed feeding studies only after first being assayed for toxins which might reasonably be expected, using chemical, biochemical or biological methods. For example, all test material from fungal sources should be assayed for certain known mycotoxins (*4, 38*).

4. In establishing an Acceptable Daily Intake for microbial enzymes based on the animal feeding studies which we have proposed, there should be no adverse effect at a dose which is 100 times the estimated mean human exposure (based on TOS). This criterion applies to the single oral challenge and to the subchronic feeding study, and is based on the traditional 100-to-1 safety factor for food chemicals (*26*).
5. The only test materials which can pass through the Decision Tree without a subchronic feeding study are those which satisfy the criteria of B.1 or C.1. i.e., certain bacteria, yeast and molds, which are well-known and have never been associated with foodborne illness or disease. However, as stated above, bacteria and yeast that meet these criteria still must pass the single oral challenge test, and molds must give negative test results for a battery of known mycotoxins.

Nonmicrobial enzymes. As indicated previously, meat animals, e.g., cattle, swine and sheep, and edible and nontoxic plants, e.g., papaya, pineapple, barley and fig, have long histories as sources of enzymes used in food processing (*3, 4, 16, 45, 47*). These traditional sources need not be subjected to toxicity testing.

For the purposes of this paper, it is assumed that only animals commonly regarded as food will be employed in enzyme manufacture. As long as CGMPs are followed during manufacture, enzymes derived from food animals may be assumed to be safe for use in food processing. Animal testing for possible toxicity is not warranted.

With regard to new plant enzyme sources, it is assumed that only edible plants will be considered. If the edible plant has been well-studied, is widely consumed without apparent harm, and does not produce toxic substances, then no animal testing should be required. However, if the plant is known to produce toxins, then care should be taken not to concentrate the toxic substances during enzyme manufacture. The final enzyme preparation should not contain toxic substances in quantities that might represent a hazard to health.

ACKNOWLEDGMENTS

We thank the members of the Ad Hoc Enzyme Technical Committee for guidance in preparation of the manuscript. This work was supported in part by the Ad Hoc Enzyme Technical Committee and by the College of Agricultural and Life Sciences, University of Wisconsin - Madison, Madison, Wisconsin.

REFERENCES

1. Ames, B. N., J. McCann, and E. Yamasaki. 1975. Methods for detecting carcinogens and mutagens with the salmonella/mammalian-microsome mutagenicity test. Mutat. Res. 31:347-364.
2. Anonymous. 1982. U.S.A. comment to the ad hoc working group on food additive intake. 15th Session Codex Committee of Food Additives, Joint FAO/WHO Food Standards Programme, Hague, Netherlands.
3. Anonymous. 1981. Food chemicals codex, 3rd ed. National academy of Science/National Research Council, Food and Nutrition Board, Committee on Codex Specifications, National Academy Press, Washington, DC. pp. 107-110, 480-481.
4. Anonymous. 1981. Specification for identity and purity of carrier solvents, emulsifiers and stabilizers, enzyme preparations, flavoring agents, food colours, sweetening agents and other food additives. Food and nutrition paper no. 19. 25th session of the Joint Expert Committee on Food Additives, FAO/WHO, Rome. pp. 215-218.
5. Anonymous. 1981. The 1978 enzyme survey, summarized data. National Academy of Sciences/National Research Council, Food and Nutrition Board, Committee on GRAS List Survey, Phase III. National Academy Press, Washington, DC.
6. Anonymous. 1981. CSPI asks hearing on proposal to rescind alcohol ingredient labeling. Food Chem. News 23(23):29-30.
7. Anonymous. 1982. Code of federal regulations. 21 CFR 173.110, 120, 130, 135, 145, 150, 160, 165. Office of the Federal Register, National Archives and Records Service, General Services Administration, Washington, DC.
8. Anonymous. 1981. Code of federal regulations. 21 CFR Sec. 110.1. Office of the Federal Register, National Archives and Records Service, General Services Administration, Washington, DC.
9. Anonymous. 1979. Modern urine chemistry. Ames Division, Miles Laboratory Inc., Elkart, IN.
10. Anonymous. 1978. Evaluation of certain food additives. pp. 8-9. *In* Evaluation of certain food additives, Technical report series no. 617. World Health Organization, Geneva.
11. Anonymous. 1977. Evaluation of the health aspects of papain as a food ingredient. Select committee on GRAS substances. National Technical Information Service Document PB-274-174. U.S. Dept. of Commerce, Springfield, VA.
12. Anonymous. 1975. Toxicological evaluation of some food colours, enzymes, flavor enhancers, thickening agents, and certain other food additives. pp. 11, 124-141. *In* World Health Organization Food Additive series, No. 6. World Health Organization, Geneva.
13. Anonymous. 1973. Toxicants occurring naturally in foods. National Academy of Science, National Research Council, Food and Nutrition Board, Committee on Food Protection, Washington, DC.
14. Anonymous. 1972. A comprehensive survey of industry on the use of food chemicals generally recognized as safe (GRAS). Comprehensive GRAS survey. National Academy of Sciences/National Research Council. National Technical Information Service PB-221 - 949. U.S. Dept. of Commerce, Springfield, VA.
15. Anonymous. 1971. Enzyme-containing laundering compounds and consumer health. National Academy of Science, National Research Council, Division of Medical Sciences, Ad Hoc Committee on Enzyme Council, Division of Medical Sciences, Ad Hoc Committee on Enzyme Detergents. National Technical Information Service, Document PB-204-118. U.S. Dept. of commerce, Springfield, VA.
16. Beckhorn, E. J., M. D. Labbe, and L. A. Underkofler. 1965. Production and use of microbial enzymes for food processing. J. Agr. Food Chem. 13:30-34.
17. Buchanan, R. E., and N. E. Gibbons (eds.). 1974. Bergey's manual of determinative bacteriology, 8th ed. Williams and Wilkins Company, Baltimore, MD.
18. Clarke, P. H. 1978. Experiments in microbial evolution. pp. 137-218. *In* I. C. Gunsalus, L. N. Ornston and J. R. Sokatch (eds.) The bacteria, vol. VI. Academic Press, New York.
19. Cooney, D. A., G. Stergis, and H. N. Jayaram. 1980. Enzymes, therapeutic. pp. 225-240. *In* Kirk-Othmer encyclopedia of chemical technology, 3rd ed., vol. 9. John Wiley & Sons, New York.
20. Dixon, R. L. 1980. Toxic responses of the reproductive system. pp. 332-354. *In* J. Doull, C. D. Klaassen and M. O. Amdur (eds.) Toxicology, 2nd ed. Macmillan Publishing Co., New York.
21. Doolittle, R. F. 1981. Similar amino acid sequences: chance or common ancestry? Science 214:149-159.
22. Ehrlich, K., A. J. DeLucca, and A. Ciegler. 1981. Isolation of toxic metobolites from *Aspergillus niger* recovered from stored cottonseed. Abstracts of the 95th Annual meeting of the Association of Official Analytical Chemists, Washington, DC. p. 67 (Abstract No. 195).
23. Fennel, D. I. *Aspergillus oryzae* (NRRL Strain 1988): a clarification. Science 194:1188.
24. Food Safety Council. 1978. Proposed system for food safety assessment. Food Cosmet. Toxicol. 16 (Suppl.):1-136.
25. Free, A. H., and H. M. Free. 1978. Urodynamics. Ames Division, Miles Laboratory, Inc., Elkhart, IN.
25a. Goldberg, I. H., T. Hatayama, L. S. Kappen, M. A. Napier, and L. F. Povrick. 1981. Protein antibiotics as DNA-damaging agents. pp. 163-191. *In* A. C. Sartorelli, J. S. Lazo, and J. R. Bertino (eds.) Molecular actions and targets for cancer chemotherapeutic agents. Academic Press, New York.
26. Hall, R. L. 1977. Safe at the plate. Nutrition Today 12 (6):6-9, 28-31.
27. Harbison, R. D. 1980. Teratogens. pp. 158-175. *In* J. Doull, D. C. Klaassen, and M. O. Amdur (eds.) Toxicology, 2nd ed. Macmillan Publishing Co., New York.
28. Hartshorn, E. A. 1973. Handbook of drug interactions, 2nd ed. Hamilton Press, Hamilton, IL.
29. Keyes, M. H. 1980. Enzymes, immobilized. pp. 148-172. *In* Kirk-Othmer encyclopedia of chemical technology, 3rd ed., vol. 9. John Wiley & Sons, New York.
30. Koch, A. L. 1981. Evolution of antibiotic resistance gene function. Microbiol. Rev. 45:355-378.
31. LeBlanc, J. R., Jr., S. Madharan, and R. E. Porter. 1978. Ammonia. pp. 470-516. *In* Kirk-Othmer encyclopedia of chemical technology, 3rd ed., vol. 2. John Wiley & Sons, New York.
32. MacAllister, R. V., E. K. Wardrip, and B. J. Schnyder. 1975. Mod-

ified starches, corn syrups containing glucose and maltose, corn syrups containing glucose and fructose, and crystalline dextrose. pp. 331-359. *In* Enzymes in food processing, 2nd ed. Academic Press, New York.

33. Miller, E. C., and J. A. Miller. 1979. Overview on the revelance of naturally occurring carcinogens, promoters, and modulators of carcinogenesis in human cancer. pp. 1-17. *In* E. C. Miller, J. A. Miller, I. Hirono, T. Subimura and S. Takayama (eds.) Naturally occurring carcinogens-mutagens and modulators of carcinogensis. University Park Press, Baltimore.
34. Miller, J. A. 1970. Carcinogenesis by chemicals: an overview. Cancer Res. 30:559-576.
35. Mills, A. L., and M. Alexander. 1976. N-nitrosamine formation by cultures of several microorganisms. Appl. Environ. Microbiol. 31:892-895.
36. Moskowitz, G. J., and T. Cayle. 1974. A method for detection of beta-nitropropionic acid in crude biological extracts. Cereal Chem. 51:96-105.
37. Ough, C. S. 1976. Ethyl carbamate in fermented beverages and foods. I. Naturally-occurring ethyl carbamate. Agr. Food Chem. 24:323-328.
38. Patterson, D. S., and B. A. Roberts. 1979. Mycotoxins in animal feedstuffs: sensitive thin layer chromatographic detection of aflatoxin, ochratoxin A, sterigmatocystin, zearalenone, and T-2 toxin. J. Assoc. Off. Anal. Chem. 62:1265-1267.
39. Pitot, H. C. 1978. Fundamentals of oncology. Marcel Dekker, New York.

39a. Povirk, L. F., and I. H. Goldberg. 1982. Covalent adducts of DNA and the nonprotein chromophore of neocarzinostatin contain a modified deoxyribose. Proc. Natl. Acad. Sci. (USA) 79:369-373.

40. Reanney, D. 1976. Extrachromosomal elements as possible agents of adaption and development. Bacteriol. Rev. 40:552-590.
41. Riemann, H., and F. L. Bryan (eds.). 1979. Food-borne infections and intoxications, 2nd ed. Academic Press, New York.
42. Rodricks, J. V., C. W. Hesseltine, and M. A. Mehlman (eds.). 1977. Mycotoxins in human and animal health. Pathotox Publishers, Park Forest South, IL.
43. Rogers, R. W. 1977. Monograph on microbial enzymes. National Technical Information Service Document PB-289-661. U.S. Dept. of Commerce, Springfield, VA.
44. Rogers, R. W. 1978. Monograph on nonmicrobial proteolytic enzymes exclusive of papain. National Technical Information Service, Document PB-289-397. U.S. Dept. of Commerce, Springfield, VA.
45. Roland, J. F. 1981. Regulation of food enzymes. Enzyme Microb. Technol. 3:105-110.
46. Rothchild, L. (ed.). 1982. FCN/PTCN annual yearbook. Food Chemical News and Pesticide & Toxic Chemical News, Washington, DC. pp.173-224.
47. Scott, D. 1980. Enzymes, industrial. pp. 173-224. In Kirk-Othmer encyclopedia of chemical technology, 3rd ed., vol. 9. John Wiley & Sons, New York.
48. Speyer, J. F. 1965. Mutagenic DNA polymerase. Biochem. Biophys. Res. Commun. 21:6-8.
49. Starace, D., and H. C. Barfoed. 1980. Enzyme detergents. pp. 138-148. *In* Kirk-Othmer encyclopedia of chemical technology, 3rd ed., vol. 9. John Wiley & Sons, New York.
50. Stoloff, L., P. Mislivec, and A. F. Schindler. 1977. *Aspergillus oryzae* (NRRL Strain 1988). Science 196:1354.
51. Thilly, W. G., and H. L. Liber. 1980. Genetic toxicology. pp. 139-157. *In* J. Doull, C. D. Klassen and M. O. Amdur (eds.) Genetic toxociology, 2nd ed. Macmillan Publishing Co., New York.
52. White, A., P. Handler, and E. L. Smith. 1973. Principles of biochemistry, 5th ed. McGraw-Hill Book Company, New York.
53. Zioudrou, D., R. A. Streaty, and W. A. Klee. 1979. Opioid peptides derived from food proteins, the exorphins. J. Biol. Chem. 254:2446-2449.

News and Events

9th Annual Food Microbiology Research Conference

The 9th Annual Food Microbiology Research Conference will be held November 2-4, 1983 in Chicago, Illinois. For more information contact: Dr. J. M. Goepfert, Canada Packers, Ltd., 2211 St. Clair Avenue West, Toronto, Ontario, Canada M6N 1K4.

Classified Ads

For Sale

Single Service milk sample tubes. For further information and a catalogue please write, Dairy Technology Inc., PO Box 101, Eugene, OR 97401.

Regulatory Toxicology and Pharmacology **33,** 173–186 (2001)
doi:10.1006/rtph.2001.1466, available online at http://www.idealibrary.com on IDEAL®

Evaluating the Safety of Microbial Enzyme Preparations Used in Food Processing: Update for a New Century[1]

Michael W. Pariza[2] and Eric A. Johnson

Food Research Institute, Department of Food Microbiology and Toxicology, University of Wisconsin–Madison, Madison, Wisconsin 53706

Received October 30, 2000; published online April 12, 2001

Microbial enzymes used in food processing are typically sold as *enzyme preparations* that contain not only a desired enzyme activity but also other metabolites of the production strain, as well as added materials such as preservatives and stabilizers. The added materials must be food grade and meet applicable regulatory standards. The purpose of this report is to present guidelines that can be used to evaluate the safety of the metabolites of the production strain that are also present in the enzyme preparation, including of course, but not limited to, the desired enzyme activity itself. This discussion builds on previously published decision tree mechanisms and includes consideration of new genetic modification technologies, for example, modifying the primary structure of enzymes to enhance specific properties that are commercially useful. The safety of the production strain remains the primary consideration in evaluating enzyme safety, in particular, the *toxigenic potential* of the production strain. Thoroughly characterized nonpathogenic, nontoxigenic microbial strains, particularly those with a history of safe use in food enzyme manufacture, are logical candidates for generating a *safe strain lineage*, through which improved strains may be derived via genetic modification by using either traditional/classical or rDNA strain improvement strategies. The elements needed to establish a safe strain lineage include thoroughly characterizing the host organism, determining the safety of all new DNA that has been introduced into the host organism, and ensuring that the procedure(s) that have been used to modify the host organism are appropriate for food use. Enzyme function may be changed by intentionally altering the amino acid sequence (e.g., protein engineering). It may be asked if such modifications might also affect the safety of an otherwise safe enzyme. We consider this question in light of what is known about the natural variation in enzyme structure and function and conclude that it is unlikely that changes which improve upon desired enzyme function will result in the creation of a toxic protein. It is prudent to assess such very small theoretical risks by conducting limited toxicological tests on engineered enzymes. The centerpiece of this report is a decision tree mechanism that updates previous enzyme safety evaluation mechanisms to accommodate advances in enzymology. We have concluded that separate mutagenicity testing is not needed if this decision tree is used to evaluate enzyme safety. Under the criteria of the decision tree, no new food enzyme can enter the market without critical evaluation of its safety.

INTRODUCTION

Microbial enzymes used in food processing are typically sold as *enzyme preparations* that contain not only a desired enzyme activity but also other metabolites of the production strain, as well as added materials such as preservatives and stabilizers. The added materials must be food grade and meet the standards of regulatory policy where the enzyme is used. The purpose of this report is to present guidelines that can be used to evaluate the safety of the metabolites of the production strain that are also present in the enzyme preparation, including of course, but not limited to, the desired enzyme activity itself. This discussion builds on previous reports (Pariza and Foster, 1983; IFBC, 1990; Kessler *et al.*, 1992) and includes consideration of new genetic modification technologies, for example, modifying the primary structure of enzymes to enhance specific properties that are commercially useful.

Many of the enzymes that were used or are currently used in food processing are listed in Table 1; enzymes that were listed in the Pariza and Foster (1983) publication are marked. Over time, new enzymes will be added to this list and some now on the list may be deleted. The Enzyme Technical Association (ETA) periodically updates this list and maintains it on their web site, http://www.enzymetechnicalassoc.org/.

[1] The authors thank the Enzyme Technical Association for their support in preparing the manuscript.

[2] To whom correspondence should be addressed. Fax: (608) 263-1114. E-mail: mwpariza@facstaff.wisc.edu.

0273-2300/01 $35.00

TABLE 1
Enzymes Used in Food Processing Today

Trivial name	Classification	Source	Systematic Names (IUB)[a]	IUB No.[a]	CAS No.[b]
α-Amylase	Carbohydrase	(1) *Aspergillus niger* var.[c] (2) *Aspergillus oryzae* var.[c] (3) *Rhizopus oryzae* var.[c] (4) *Bacillus subtilis* var.[c] (5) *Bacillus amyloliquefaciens*[d] d-*Bacillus* (6) Barley malt[c] (7) *Bacillus licheniformis*[d] d-*Bacillus licheniformis* (8) *Bacillus stearothermophilus* (9) *Bacillus subtilis*[d] d-*Bacillus megaterium* (10) *Bacillus subtilis*[d] d-*Bacillus stearothermophilus* (11) *Microbacterium imperiale* (12) *Bacillus amyloliquefaciens* (13) *Bacillus licheniformis* var.[c] (14) *Bacillus licheniformis*[d] d-*Bacillus stearothermophilus* (15) *Aspergillus niger*[d] d-*Aspergillus niger*	1,4-α-D-Glucan glucanohydrolase	3.2.1.1	9000-90-2
Aminopeptidase	Protease	(1) *Aspergillus niger* (2) *Rhizopus oryzae* (3) *Aspergillus oryzae*	α-Aminoacyl-peptide hydrolase	3.4.11.11	
AMP deaminase	Adenosine deaminase	*Aspergillus melleus*	AMP aminohydrolase	3.5.4.6	9025-10-9
Arabinofuranosidase	Carbohydrase	(1) *Aspergillus niger* (2) *Aspergillus niger*[d] d-*Aspergillus niger*		3.2.1.55	9067-74-7
β-Amylase	Carbohydrase	(1) Barley malt[c] (2) Barley (ungerminated)	1,4-α-D-Glucan maltohydrolase	3.2.1.2	9000-91-3
Bromelain	Protease	Pineapples: *Ananas comosus* *Ananas bracteatus* (L)[c]	None	3.4.22.32 3.4.22.33	37189-34-7 9001-00-7
Catalase	Oxidoreductase	(1) *Aspergillus niger*[d] d-*Aspergillus niger* (2) bovine liver[c] (3) *Micrococcus luteus*[c] (4) *Aspergillus niger* var.[c]	Hydrogen peroxide: hydrogen peroxide oxidoreductase	1.11.1.6	9001-05-2
Cellulase	Carbohydrase	(1) *Aspergillus niger* var.[c] (2) *Trichoderma reesei*[c] (formerly *longibrachiatum*) (3) *Trichoderma reesei*[d] d-*Trichoderma reesei* (4) *Trichoderma viride* (5) *Aspergillus aculeatus*	Endo-1,4-(1,3; 1,4)-β-D-glucan 4-glucanohydrolase	3.2.1.4	9012-54-8
Chymosin	Protease	(1) *Aspergillus niger* var. *awamori*[d] d-calf prochymosin gene (2) *Escherichia coli* K-12[d] d-calf prochymosin gene (3) *Kluyveromyces marxianus*[d] d-calf prochymosin gene	Cleaves a single bond in κ-casein	3.4.23.4	9001-98-3
Chymotrypsin	Protease	Bovine or porcine pancreatic extract	None	3.4.21.1	9004-07-3
Dextranase	Carbohydrase	(1) *Chaetomium erraticum* (2) *Chaetomium gracile*	1,6-α-D-Glucan 6-glucanohydrolase	3.2.1.11	9025-70-1
Esterase	Lipase	*Rhizomucor miehei*		3.1.1.3	9001-62-1
Ficin	Protease	Figs: *Ficus* sp.[c]	None	3.4.22.3	9001-33-6
α-Galactosidase	Carbohydrase	(1) *Mortierella vinacea* var. *raffinoseutilizer* (2) *Aspergillus niger* (3) *Saccharomyces cerevisiae*[d] d-Guar seed	α-D-Galactoside galactohydrolase	3.2.1.22	90025-35-8

TABLE 1—*Continued*

Trivial name	Classification	Source	Systematic Names IUB[a]	IUB No[a]	CAS No[b]
β-Glucanase	Carbohydrase	(1) *Aspergillus niger* var.[c] (2) *Bacillus subtilis* var.[c] (3) *Trichoderma reesei* (formerly *longibrachiatum*) (4) *Talaromyces emersonii* (formerly *Penicillium emersonii*) (5) *Bacillus amyloliquefaciens* (6) *Aspergillus aculeatus* (7) *Bacillus amyloliquefaciens*[d] d-*Bacillus amyloliquefaciens* (8) *Disporotrichum dimorphosporum*	1,3-(1,3; 1,4)-β-D-Glucan 3(4)-glucanohydrolase	3.2.1.6	62213-14-3
Glucoamylase (amyloglucosidase)	Carbohydrase	(1) *Aspergillus niger* var.[c] (2) *Aspergillus oryzae* var.[c] (3) *Rhizopus oryzae* var.[c] (4) *Rhizopus niveus* (5) *Rhizopus delemar* (6) *Aspergillus niger*[d] d-*Aspergillus niger*	1-4-α-D-Glucan glucohydrolase	3.2.1.3	9032-08-0
Glucose isomerase	Isomerase	(1) *Actinoplanes missouriensis*[c] (2) *Bacillus coagulans*[c] (3) *Streptomyces olivaceus*[c] (4) *Streptomyces olivochromogenes*[c] (5) *Streptomyces rubiginosus*[d] d-*Streptomyces rubiginosus* (6) *Streptomyces murinus* (7) *Microbacterium arborescens* (8) *Streptomyces rubiginosus*[c]	D-Xylose ketoisomerase	5.3.1.5	9055-00-9
Glucose oxidase	Oxidoreductase	(1) *Aspergillus niger*[d] d-*Aspergillus niger* (2) *Aspergillus niger* var.[c]	β-D-Glucose: oxygen 1-oxidoreductase	1.1.3.4	9001-37-0
Glutaminase	Glutaminase	*Bacillus subtilis*	L-Glutamate aminohydrolase	3.5.1.2	9001-47-2
β-D-Glucosidase	Carbohydrase	(1) *Aspergillus niger* var. (2) *Trichoderma reesei* (formerly *longibrachiatum*)	β-D-Glucoside glucohydrolase	3.2.1.21	9001-22-3
Hemicellulase[e]	Carbohydrase	(1) *Aspergillus niger* var.[c] (2) *Aspergillus aculeatus* (3) *Aspergillus foetidus*	(1) α-L-Arabinofuranoside arabinofuranohydrolase (2) 1,4-β-D-Mannan mannanohydrolase (3) 1,3-β-D-Xylan-xylanohydrolase (4) 1,5-α-L-Arabinan 1,5-α-L arabinanohydrolase (5) 1,4,β-D-Xylan xylanohydrolase (6) 1,4,β-D-Xylan xylohydrolase (7) Endo-1,4-β-D-xylanase	3.2.1.55 3.2.1.78 3.2.1.32 3.2.1.99 3.2.1.8 3.2.1.37	9025-56-3 9025-57-4
Hesperidinase	Carbohydrase	*Penicillium decumbens*	α-L-Rhamnoside rhamnohydrolase	3.2.1.40	37288-35-0
Invertase	Carbohydrase	*Saccharomyces* sp. (*Kluyveromyces*)[c]	β-D-Fructofuranoside fructohydrolase	3.2.1.26	9001-57-4
Lactase	Carbohydrase	(1) *Aspergillus niger* var. [c] (2) *Aspergillus oryzae* var.[c] (3) *Saccharomyces* sp.[c] (4) *Candida pseudotropicalis* (5) *Kluyveromyces marxianus* var. *lactis* (6) *Kluyveromyces marxianus* var. *lactis*[d] d-*Kluyveromyces marxianus* var. *lactis* (7) *Aspergillus oryzae*[d] d-*Aspergillus oryzae*	β-D-Galactoside galactohydrolase	3.2.1.23	9031-11-2

TABLE 1—*Continued*

Trivial name	Classification	Source	Systematic Names IUB[a]	IUB No[a]	CAS No[b]
Lipase	Lipase	(1) Edible forestomach tissue of calves, kids, and lambs[c] (2) Animal pancreatic tissues[c] (3) *Aspergillus oryzae* var.[c] (4) *Aspergillus niger* var.[c] (5) *Rhizomucor miehei* (6) *Candida rugosa* (7) *Candida lipolytica* (8) *Rhizopus delemar* (9) *Rhizopus oryzae* (10) *Rhizopus niveus* (11) *Penicillium roqueforti* (12) *Penicillium camembertii* (13) *Mucor javanicus* (14) *Aspergillus oryzae*[d] d-*Rhizomucor miehei* (15) *Aspergillus oryzae*[d] d-*Thermomyces lanuginosus*		3.1.1.3	9001-62-1
Maltogenic amylase	Carbohydrase	*Bacillus subtilis*[d] d-*Bacillus stearothermophilus*	1,4-α-D-Glucan α-maltohydrolase	3.2.1.133	160611-47-2
Naringinase	Carbohydrase	*Penicillium decumbens*	α-L-Rhamnoside rhamnohydrolase	3.2.1.40	37288-35-0
Pancreatin	Mixed: carbohydrase, lipase, and protease	Bovine and porcine pancreatic tissue	(1) 1,4-α-D-Glucan glucanohydrolase (2) Triacylglycerol acylhydrolase (3) Protease	3.2.1.1 3.1.1.3 3.4.21.4	9000-90-2 9001-62-1 9002-07-7
Papain	Protease	Papaya: *Carica papaya (L)*[c]	None	3.4.22.2	9001-73-4
Pectin esterase	Carbohydrase	(1) *Aspergillus oryzae*[d] d-*Aspergillus aculeatus* (2) *Aspergillus niger* var. (3) *Aspergillus niger*[d] d-*Aspergillus niger*	Pectin pectylhydrolase	3.1.1.11	9025-98-3
Pectin lyase	Carbohydrase	*Aspergillus niger*		4.2.2.10	90025-98-3
Pectinase[e]	Carbohydrase	(1) *Aspergillus niger* var.[c] (2) *Rhizopus oryzae* var.[c] (3) *Aspergillus aculeatus*	(1) Poly(1,4-α-D-galacturonide) glycanohydrolase (2) Pectin pectylhydrolase (3) Poly(1,4-α-D-galacturonide) lyase (4) Pectin lyase (5) L-Arabinofuranoside arabinofuranohydrolase (6) 1,5-L-Arabinan arabinofuranohydrolase (7) Exo-polygalacturonase (8) Endo-1,4-β-galactanase (9) Pectin acetylesterase (10) Exopolygalacturonase lyase	3.2.1.15 3.1.1.11 4.2.2.2 4.2.2.10 3.2.1.55 3.2.1.99 3.2.1.67 3.2.1.89 3.1.1.6 4.2.2.9	9032-75-1 9025-98-3 9015-75-2 9033-35-6 9067-74-7 37325-54-5
Pepsin	Protease	Porcine or other animal stomach tissue[c]	None	3.4.23.1 3.4.23.2	9001-75-6 9025-48-3
Phosphodiesterase	Nuclease	(1) *Penicillium citrinum* (2) *Leptographium procerum* (formerly *Verticicladiella procera)*	Oligonucleate 5′-nucleotidohydrolase	3.1.4.1	9025-82-5
Phospholipase A_2	Lipase	(1) *Animal pancreatic tissue* (2) *Streptomyces violaceoruber* (3) *Aspergillus niger*[d] d-porcine *pancreas*	(1) Phosphatidylcholine 2-acylhydrolase	3.1.1.4	9001-84-7
Phytase	Phosphatase	*Aspergillus niger*[d] d-*Aspergillus niger*	(1) Myo-inositol-hexakisphosphate--3-phosphohydrolase (2) Orthophosphoric monoester phosphohydrolase	3.1.3.8 3.1.3.2	37288-11-2 9001-77-8

TABLE 1—*Continued*

Trivial name	Classification	Source	Systematic Names IUB[a]	IUB No[a]	CAS No[b]
Protease (general)	Protease	(1) *Aspergillus niger* var.[c]	None	3.4.23.18	9025-49-4
		(2) *Aspergillus oryzae* var.[c]			
		(3) *Aspergillus melleus*		3.4.21.14	9014-01-1
		(4) *Bacillus subtilis*		3.4.21.62	9014-01-1
		(5) *Bacillus subtilis*[d] d-*Bacillus amyloliquefaciens*		3.4.24.28	76774-43-1
		(6) *Bacilus amyloliquefaciens*[d] d-*Bacillus amyloliquefaciens*			9068-59-1 9073-79-4
		(7) *Bacillus licheniformis* var.		3.4.24.4	9001-61-0
		(8) *Bacillus stearothermophilus*		3.4.23.6	9080-56-2
		(9) *Rhizopus niveus*		3.4.11.1	
		(10) *Rhizopus oryzae*			
		(11) *Bacillus amyloliquefaciens*			
		(12) *Aspergillus oryzae*[d] d-*Rhizomucor miehei*			
Pullulanase	Carbohydrase	(1) *Bacillus acidopullulyticus*	α-Dextrin 6-glucanohydrolase	3.2.1.41	9075-68-7
		(2) *Bacillus licheniformis*[d] d-*Bacillus deramificans*			
		(3) *Bacillus naganoensis*[d]			
		(4) *Bacillus subtilis*[d] d-*Bacillus naganoensis*			
		(5) *Bacillus circulans*	Pullulan 6-glucanohydrolase		
Rennet	Protease	(1) Fourth stomach of ruminant animals	None	3.4.23.4	9001-98-3
				3.4.23.22	37205-60-0
		(2) *Endothia parasitica*		3.4.23.23	148465-73-0
		(3) *Rhizomucor miehei*			
		(4) *Rhizomucor pusillus* (Lindt)			
		(5) *Aspergillus oryzae*[d] d-*Rhizomucor miehei*			
Transglucosidase	Glucanotransferase	*Aspergillus niger*	1,4-α-D-Glucan 4-α-D-glycosyltransferase	2.4.1.25	9032-09-1
Transglutaminase	Acyltransferase or aminotransferase	*Streptoverticillium mobaraense* var.	*R*-Glutaminly-peptide: amine γ-glutamyltransferase	2.3.2.13	80146-85-6
Trypsin	Protease	Animal pancreas	None	3.4.21.4	9002-07-7
Urease	Protease	*Lactobacillus fermentium*	None	3.5.1.5	9002-13-5
Xylanase	Carbohydrase	(1) *Trichoderma longibrachiatum*[d] d-*Trichoderma longibrachiatum*	(1) 1,4-β-D-Xylan xylanohydrolase	3.2.1.8	9025-57-4
			(2) 1,3 β-D-Xylan xylanohydrolase	3.2.1.32	9025-55-2
		(2) *Aspergillus niger var. awamori*[d] d-*Aspergillus* var.	(3) Endo-1,4(3)-β-D-hemicellulase		
		(3) *Bacillus licheniformis*[d] d-*Bacillus licheniformis*			
		(4) *Aspergillus oryzae*[d] d-*Thermomyces lanuginosus*			
		(5) *Disporotrichum dimorphosporum*			
		(6) *Aspergillus niger*[d] d-*Aspergillus niger*			
		(7) *Trichoderma reesei* (formerly *longibrachiatum)*			
		(8) *Bacillus subtilis*[d] d-*Bacillus subtilis*			

[a] Enzyme nomenclature primarily based on the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB). See also http://www.expasy.ch/enzyme/.
[b] *Chemical Abstract Service Registry Number*.
[c] Originally listed in the Pariza and Foster (1983) publication.
[d] A genetically modified organism. The donor organism is listed after "d-."
[e] Usually a mixture of the activities listed under the systematic name.

It is instructive to compare the number of enzymes from the 1983 list and today because the number of enzymes, and the microbial species from which production strains are derived, have greatly expanded in the past 20 years. This has occurred in response to the constantly evolving requirements of a very diverse international food processing industry.

For example, no enzyme listed in the 1983 publication is a product of recombinant DNA (rDNA) technology, because these methods were not yet in commercial use in

1983. Of course this changed in the ensuing years, and many of the production strains listed in Table 1 have been improved using rDNA technology. However, and most importantly, each of the enzymes and production strains listed in Table 1 has been evaluated for safety using criteria that are comparable to those described by Pariza and Foster (1983) and IFBC (1990).

CONSIDERATIONS PERTAINING TO FOOD ENZYME SAFETY EVALUATION

Safety of the Production Strain

Pariza and Foster (1983) discussed safety considerations for food-processing enzymes derived from production strains that were improved via traditional (nonrecombinant) methodologies. The following issues were considered: the safety of the production strain (referred to as the "source organism") with particular regard to toxigenic and pathogenic potential; allergies and primary irritations; carcinogens and mutagens; teratogens and reproductive effects; antibiotics; products of enzymatic reactions; interactions between enzymes and other food components; and direct effects of food enzymes on consumers.

It was concluded that the safety of the production strain should be the primary consideration in evaluating enzyme safety. The primary issue in evaluating the safety of a production strain is its *toxigenic potential*, specifically the possible synthesis by the production strain of toxins that are active via the oral route. *Pathogenic potential* is not usually an area of concern for consumer safety because enzyme preparations rarely contain viable organisms. Pathogenicity is, however, important to worker safety.

Toxigenic potential. Microbial toxins that are active via the oral route may be produced by certain bacteria or certain filamentous fungi (molds). Yeasts, by contrast, are not known to produce such toxins.

The oral toxins produced by bacteria cause food poisoning. They are proteinaceous in nature and elicit a rapid response. *In vivo* and/or *in vitro* tests are available for the detection of these toxins. The principal bacterial food poisoning toxins have been purified and many of their corresponding genes have been sequenced. Most of the toxins have been well characterized and exhaustively studied (Aktories and Just, 2000; Alouf and Freer, 1999; Rappuoli and Montecucco, 1997). The bacteria and fungi that produce these toxins have also been extensively characterized (Doyle *et al.*, 1997; Fischetti, 2000). This information provides the basis for testing new bacterial isolates for toxigenic potential.

The oral toxins produced by filamentous fungi are small molecular weight organic molecules, usually less than 1000 Da in size (Chu, 2000). These are referred to as mycotoxins. Most mycotoxins are acutely toxic, and many of them may also induce chronic toxicity (e.g., cancer) and developmental toxic effects when repeatedly administered to test animals. Chemical tests have been developed for the more important known mycotoxins (Chu, 2000). These tests can measure low levels of mycotoxins that would not elicit an acute response.

Pathogenic potential. It is extremely unlikely that a frank human pathogen would ever be used in food enzyme manufacture. Moreover, food enzyme preparations rarely contain viable production organisms. Hence the issue of pathogenicity is largely moot as regards food enzyme production strains. Nonetheless it is common industrial practice to evaluate previously untested host organisms for potential pathogenicity, using animal models.

It is important to distinguish between pathogenicity and opportunistic infection. Many microorganisms will produce opportunistic infections if they gain access to tissue sites that are normally protected by host barriers. Examples are infections in deep wounds produced by otherwise harmless microbes or infections by normally harmless microorganisms in individuals with a compromised immune system. By contrast, a true pathogen will produce disease or infection in an individual who would otherwise be considered healthy. Hence, a pathogen must be able to cross or evade noncompromised host barriers (Falkow, 1997; Mims, 1991). Accordingly, one cannot assess potential pathogenicity in compromised hosts.

It is important not to confuse the effect of the microbe itself with the host response to the microbe. For example, injecting dead bacteria into animals may lead to a catabolic cascade that may end in death by septic shock. This is not the result of pathogenicity, since the bacteria are dead and cannot produce an infection. Rather the reaction is caused by the release of hormone-like substances (*cytokines*) from the host immune cells in response to the presence of the dead bacteria (Beutler and Cerami, 1997). Accordingly, simply injecting microorganisms into animals is not an appropriate way to assess potential pathogenicity.

Information on the human pathogenic potential of microorganisms is readily available, for example, at http://www.cdc.gov/ncidod/dvbid/Biosafety_manual_rev_1994.pdf, or the NIH Guidelines for Research Involving Recombinant DNA Molecules at http://www4.od.nih.gov/oba/guidelines.html.

Safe Strain Lineage

Thoroughly characterized nonpathogenic, nontoxigenic microbial strains, particularly those with a history of safe use in food enzyme manufacture, are logical candidates for generating a *safe strain lineage*, through which improved strains may be derived via genetic modification either by using traditional/classical or rDNA strain improvement strategies (IFBC, 1990). The

elements needed to establish a safe strain lineage include thoroughly characterizing the host organism, determining the safety of all new DNA that has been introduced into the host organism, and ensuring that the procedure(s) that have been used to modify the host organism are appropriate for food use.

Historically the original isolates, from which contemporary microbial enzyme production strains were derived, were screened for vigorous growth under industrial scale fermentation conditions and selected for adequate yield of the desired enzyme product. The isolates were then subjected to mutagenesis (chemically induced or UV induced) and screened for randomly induced changes that would further increase yields. More recently rDNA technology has provided a directed means of enhancing strain improvement programs as well as for isolating and manipulating the genes coding for desired enzyme proteins.

Microbial taxonomy has advanced greatly during the past decade, particularly for bacteria and filamentous fungi (Balows and Duerden, 1998; Claridge *et al.*, 1997; Fink, 1999; Pitt and Hocking, 1997; Ward, 1998). The ability to amplify genomic DNA and obtain the sequences of microbial genomes has revolutionized our understanding of microbial taxonomy, phylogeny, and pathogenicity (Fink, 1999; Strauss and Falkow, 1997). DNA sequence data coupled with phenotypic analyses permit an accurate assessment of the taxonomy of donor and production organisms used for industrial enzyme development and manufacture. The safety evaluations of several microorganisms used in the food industry and for enzyme manufacture have been published including *Aspergillus oryzae* (Barbesgaard *et al.*, 1992), *Bacillus licheniformis* (de Boer *et al.*, 1994), *Bacillus subtilis* and *Bacillus amyloliquefaciens* (de Boer and Diderichsen, 1991), *Kluyveromyces lactis* (Bonekamp and Oosterom, 1994), and *Trichoderma reesei* (Nevalainen *et al.*, 1994). Recent technical and taxonomic treatises have provided considerable insight into the safety of enzyme production organisms based on detailed studies of phylogenetics and systematics of bacteria, yeasts, and fungi (Balows and Duerden, 1998; Kurtzman and Fell, 1999; Pitt and Hocking, 1997; Wolf, 1995). Hence, using traditional and modern molecular techniques, it is now possible to precisely determine the degree of relationships of microorganisms used in food enzyme production (e.g., Geiser *et al.*, 1998; Kuhls *et al.*, 1996). Further, according to the United States Food and Drug Administration (FDA, 1993), "... if internationally accepted rules of nomenclature are followed, changes in the taxonomic placement of an organism should not affect the ability to identify scientific references to the organism of interest, including scientific references to its toxigenicity, pathogenicity, or use in the production of food or enzymes."

Improvements in the efficiency of enzyme production may be achieved by developing microbial strains that directly or indirectly increase the amount of enzyme protein that is synthesized and available for harvest and purification (Archer and Peberdy, 1997; Demain and Davies, 1999). Genetic modifications are utilized to increase growth rate, expand the number of gene copies, enhance gene expression, and elevate enzyme secretion.

Traditional and molecular genetic techniques are also used to reduce or eliminate specific undesired endogenous enzyme activities or other characteristics. In some cases, these side activities may cause unwanted reactions in particular applications. For example, many microbial species secrete copious amounts of various proteases. Although this characteristic is desirable for certain applications, it can also lead to unwanted degradation of other secreted enzymes or have undesirable effects in certain food applications. Specific production strains have been developed in which one or several protease genes have been deleted or inactivated.

In some cases potentially useful enzyme activities have been discovered in microorganisms that are not suitable for use as production organisms in industrial fermentation. In these cases the genetic sequence encoding the desired enzyme protein can be cloned from the unsuitable microorganism and then transferred to and subsequently expressed in a well-characterized production strain with a history of safe use. This heterologous expression of enzymes is commonplace in the industry today and will remain standard practice in the commercialization of many enzymes.

Further improvements of enzyme producing microorganisms through genetic modification will continue as more knowledge is accumulated on the molecular basis of gene expression in the commonly used yeast, fungal, and bacterial enzyme production strains.

Engineered Enzymes

Protein engineering is the intentional alteration of the amino acid sequence of a protein to affect function (Arnold and Volkov, 1999; Atwell and Wells, 1999; Cleland and Craik, 1996, 1999; Kuchner and Arnold; 1997, Shaw *et al.*, 1999). This can be accomplished by inducing or introducing random mutations by chemical mutagenesis, UV irradiation, mutator strains, error-prone PCR, and other related techniques. Alternatively, site-directed mutagenesis techniques can be used to target changes to specific sites in the gene that are thought to be responsible for a particular protein characteristic. The more recent approach of directed molecular evolution employs other techniques, such as gene shuffling, to recombine gene fragments or functional blocks of gene sequence to generate variants of the protein gene sequence. In all cases, an effective and efficient selection or screening method is required to identify the altered protein having the desired functional characteristic.

When applied to enzymes, protein engineering can modify specific properties and improve the enzyme for

1. Is the production strain[a] genetically modified?[b]
   If yes,[c] go to 2. If no, go to 6.
2. Is the production strain modified using rDNA techniques?
   If yes, go to 3. If no, go to 5.
3. Issues relating to the introduced DNA[d,e] are addressed in 3a–3e.
   3a. Do the expressed enzyme product(s) which are encoded by the introduced DNA have a history of safe use in food?[f]
   If yes, go to 3c. If no, go to 3b
   3b. Is the NOAEL[g] for the test article[h] in appropriate short-term oral studies[i] sufficiently high[j] to ensure safety?
   If yes, go to 3c. If no, go to 12.
   3c. Is the test article free of transferable antibiotic resistance gene DNA?[k]
   If yes, go to 3e. If no, go to 3d.
   3d. Does the resistance gene(s) code for resistance to a drug substance used in treatment of disease agents in man or animal?
   If yes, go to 12. If no, go to 3e.
   3e. Is all other introduced DNA well characterized and free of attributes that would render it unsafe for constructing microorganisms to be used to produce food-grade products?
   If yes, go to 4. If no, go to 12.
4. Is the introduced DNA randomly integrated into the chromosome?
   If yes, go to 5. If no, go to 6.
5. Is the production strain sufficiently well characterized so that one may reasonably conclude that unintended pleiotropic effects which may result in the synthesis of toxins or other unsafe metabolites will not arise due to the genetic modification method that was employed?
   If yes, go to 6. If no, go to 7.
6. Is the production strain derived from a safe lineage, as previously demonstrated by repeated assessment via this evaluation procedure?[l]
   If yes, the test article is ACCEPTED. If no, go to 7.
7. Is the organism nonpathogenic?[m]
   If yes, go to 8. If no, go to 12.
8. Is the test article free of antibiotics?[n]
   If yes, go to 9. If no, go to 12.
9. Is the test article free of oral toxins[o] known to be produced by other members of the same species?
   If yes, go to 11. If no, go to 10.
10. Are the amounts of such toxins in the test article below levels of concern?[p]
   If yes, go to 11. If no, go to 12.
11. Is the NOAEL[q] for the test article in appropriate oral studies sufficiently high to ensure safety?
   If yes, the test article is ACCEPTED.[r]
   If no, go to 12.
12. An undesirable trait or substance may be present and the test article is not acceptable for food use. If the genetic potential for producing the undesirable trait or substance can be permanently inactivated or deleted, the test article may be passed through the decision tree again.

[a] Production strain refers to the microbial strain that will be used in enzyme manufacture. It is assumed that the production strain is nonpathogenic, nontoxigenic, and thoroughly characterized; steps 6–11 are intended to ensure this.

[b] The term "genetically modified" refers to any modification of the strain's DNA, including the use of traditional methods (e.g., UV or chemically-induced mutagenesis) or rDNA technologies.

[c] If the answer to this or any other question in the decision tree is unknown, or not determined, the answer is then considered to be NO.

[d] Introduced DNA refers to all DNA sequences introduced into the production organism, including vector and other sequences incorporated during genetic construction, DNA encoding any antibiotic resistance gene, and DNA encoding the desired enzyme product. The vector and other sequences may include selectable marker genes other than antibiotic resistance, noncoding regulatory sequences for the controlled expression of the desired enzyme product, restriction enzyme sites and/or linker sequences, intermediate host sequences, and sequences required for vector maintenance, integration, replication, and/or manipulation. These sequences may be derived wholly from naturally occurring organisms or incorporate specific nucleotide changes introduced by *in vitro* techniques, or they may be entirely synthetic.

[e] If the genetic modification served only to delete host DNA, and if no heterologous DNA remains within the organism, then proceed to step 5.

[f] Engineered enzymes are considered *not* to have a history of safe use in food, unless they are derived from a safe lineage of previously tested engineered enzymes expressed in the same host using the same modification system.

FIG. 1. Decision tree for evaluating the safety of microbially derived food enzymes.

[g] NOAEL is the acronym for the term *no observed adverse effect level*. It is the maximum dose of a test article (see below) that may be given to an animal in an appropriate repeated-dose oral toxicity test (see below) that does not produce an adverse effect. Ordinarily a NOAEL is derived from long-term feeding studies. However, given the established database indicating that microbial toxins that are active via the oral route are also acute toxins (see text), we conclude that in this very limited case, for the purposes of enzyme safety evaluation, a NOAEL can be determined from short-term gavage/feeding tests described in footnote *i*.

[h] Test article refers to the enzyme-containing material that is actually tested. It may differ from the commercial enzyme preparation in either form or formulation. For example, a lyophilized powder or other concentrated form of the enzyme preparation may be required so that it can be administered to test animals at higher concentrations. The test article may differ in formulation in that it is devoid of these safe and suitable preservatives, stabilizers, or other materials that may be commonly added to the enzyme that could affect palatability, nutrition, or some other aspect of a toxicity study.

The process for producing the test article should be representative of the process used for the final enzyme product. For instance, one would not produce a test article from a submerged culture growing in a yeast extract medium if the enzyme is to be produced commercially from cultures growing in a soy-based medium in open trays. The test article is often produced using the production process, stopping before the final purification and formulation steps.

A test article intended for pathogenicity testing would most likely consist of viable vegetative cells, spores, conidia, or other reproductive cells depending on the microorganism under test. These are commonly suspended in water, buffer, or other materials to minimize trauma to the test animals.

[i] We consider two animal toxicity tests to be appropriate for evaluating the safety of enzymes. Both are conducted using the oral route of administration as that is the intended route of exposure for consumers. The choice of which test to use is made on a case-by-case basis, depending largely on the species of the host organism and consideration of the nature of toxins that could theoretically be present (e.g., mycotoxin or bacterial enterotoxin).

The first test is an acute oral toxicity test in the rat following a single dose of the test article, as proposed by Pariza and Foster (1983). The dose to be used for this test should be at least 100 times the estimated mean human exposure (based on total organic solids (TOS)) or at least 2000 mg/kg body wt according to the OECD (Organization for Economic Cooperation and Development) Guideline for Testing of Chemicals, Acute Oral Toxicity, Guideline No. 401, Limit Test only (adopted on February 24, 1987) (Paris, 1983).

It should be emphasized that this test is not intended to establish a LD_{50}. Rather it is a specially designed test for determining the safety of bacterial enzymes, since with rare exception the only toxins known to be produced by bacteria are proteins or peptides (enterotoxins and certain neurotoxins) which are acute toxins that are produced by only a few bacterial species.

The second proposed toxicity test is a repeated-dose oral study (14–91 days) in one animal species, preferably the rat because of the historical data available on this species. The test article can be administered either in the feed or via gavage. The lowest dose used for this study should be at least 100 times the estimated mean human exposure (based on TOS). This test will detect toxicity that would be associated with the known microbial toxins that are active via the oral route.

In addition, all new enzymes should be analyzed for toxins that might be reasonably expected, using chemical, biochemical, or biological methods. For example, all test material from mold sources should be assayed for mycotoxins that are known to be synthesized by closely related species. Aflatoxins, zearalenone, T-2 toxin, ochratoxin A, and sterigmatocystin analyses are required by JECFA for all enzyme products produced by any mold (Patterson and Roberts, 1979).

[j] The NOAEL should provide at least a 100-fold margin of safety for human consumption, calculated using standard methods (Klaassen, 1996; Lehman and Fitzhugh, 1954; ILSI, 1997).

[k] Antibiotic resistance genes are commonly used in the genetic construction of enzyme production strains to identify, select, and stabilize cells carrying introduced DNA. Principles for the safe use of antibiotic resistance genes in the manufacture of food and feed products have been developed (IFBC, 1990; "FDA Guidance for Industry: Use of Antibiotic Resistance Marker Genes in Transgenic Plants," http://vm.cfsan.fda.gov/~dms/opa-armg.html).

[l] In determining safe strain lineage one should consider the host organism, all of the introduced DNA, and the methods used to genetically modify the host (see text). In some instances the procedures described by Pariza and Foster (1983) and IFBC (1990) may be considered comparable to this evaluation procedure in establishing a safe strain lineage.

[m] The issue of pathogenicity is addressed in the text. Unless an enzyme preparation contains live organisms, we do not consider this to be a relevant consumer safety issue.

[n] In this context the term antibiotic refers to antimicrobial substances that are positive in the JECFA test (FAO, 1981).

[o] The toxins of concern for food enzymes are those which are active via the oral route.

FIG. 1—*Continued*

[p] It is important to distinguish between a toxin and a toxic effect. A toxin is a chemical substance that produces a toxic effect when administered to an animal or a human in an amount that is high enough to induce the toxic effect. At lower levels of exposure there may be no adverse effect and no cause for concern (Klaassen, 1995).

[q] In the case of enzymes that lack a history of safe use in foods that are produced via genetically modified production strains, for example, newly isolated enzymes, this issue will have been addressed at step 3b. In these cases it is not necessary to repeat the animal tests again.

[r] In case of a new strain (new isolate) this testing will be fully comprehensive per this decision mechanism. However, with increased knowledge of strain and background (safe strain lineage with a number of products from this lineage tested) the depth of safety testing may be reduced. This is a case-by-case judgment; e.g. Scientific Committee for Foods (SCF) Guidelines, Section 10, also gives exemptions from the basic-full-toxicologic requirements (Scientific Committee for Food, 1992).

FIG. 1—*Continued*

particular applications. Examples include changing the pH optimum, increasing thermal stability, reducing the requirement for cofactors such as metal ions, and stabilizing the enzyme against chemical oxidation.

It may be asked if such modifications might also affect the safety of an otherwise safe enzyme. To address this question we should consider what is known about the natural variation in enzyme structure and function.

The enzymes in Table 1 are listed by their systematic names according to the Nomenclature Committee of the International Union of Biochemistry (i.e., their IUB or EC number) (IUB, 1992) and by their Chemical Abstract Service (CAS) registry number. However, in addition to the traditional IUB scheme which is based mainly on catalytic activity, several databases are available that describe the nucleotide sequences of the genes encoding the enzymes, the corresponding amino acid sequences, and information regarding the three-dimensional structures of enzymes (Brenner *et al.*, 1998; Doolittle, 1996). This information is valuable for determining the evolutionary relationships of enzymes and provides considerable insight into their structure/function determinants. Molecular analyses have demonstrated that enzymes within a given class are composed of characteristic folds that comprise domains within the entire enzyme molecule (Creighton, 1993; Doolittle, 1996; Henrissat and Davies, 1997; Jancek *et al.*, 1999).

Families/superfamilies of enzymes obtained from microorganisms found in diverse habitats retain their general tertiary structure and enzymatic properties (Siezen and Leunissen, 1997; Jancek *et al.*, 1999; Conrad *et al.*, 1995; Todd *et al.*, 1999) but they may differ in certain functional characteristics such as stability and substrate specificity (Creighton, 1993). We know of no instance in which such natural variation within enzyme families has resulted in the generation of a toxin active via the oral route. This also follows from the observation that toxicity is an unusual property among proteins. Pariza and Foster (1983) pointed out that whereas there are thousands of proteins in food, only a tiny fraction exhibit toxigenic potential by the oral route. These known toxigenic proteins differ greatly in structure from commercially available food enzymes.

Extensive studies on engineered enzymes have also demonstrated that enzymes within families/superfamilies (e.g., subtilases) that are altered by these techniques still retain their characteristic three-dimensional structure and catalytic activities (Bott *et al.*, 1992). Hence, engineered enzymes exhibit variation that is similar to that observed in nature.

An examination of enzyme structure and function indicates that it is unlikely that changes which improve upon desired enzyme function will result in the creation of a toxic protein. In our opinion it is prudent to assess this very small theoretical risk by conducting limited toxicological tests on engineered enzymes. We anticipate that when a manufacturer synthesizes a series of products through protein engineering, inserting the engineered gene into the same host with the same vector system and demonstrating through appropriate toxicological testing that each product is safe, there will come a point after which further testing of additional similar products should be considered redundant and unnecessary. The point at which this may occur would be established by independent experts on a case-by-case basis.

These conclusions should be reassessed on a regular basis, as the body of knowledge from such testing grows.

A Word about in Vitro Genotoxicity Testing

Despite the questions raised by Pariza and Foster (1983) regarding the scientific rationale and need for testing new food enzyme preparations for mutagenic activity, the practice continues, driven largely by regulatory requirements in some locales. Accordingly it is worth noting that, to our knowledge, the requirement that new enzyme preparations be tested for *in vitro* genotoxicity has failed to reveal the presence of a single mutagen or clastogen that would not have been detected using the more comprehensive decision tree approach described by Pariza and Foster (1983) and IFBC (1990), which involve analytical chemistry and limited animal feeding tests.

There appear to be three reasons for this, as follows. Proteins, including food-borne enterotoxins and neurotoxins produced by some bacteria, are not genotoxic. All known mycotoxins, some of which are genotoxic, also induce other toxic effects in test animals that are easily determined in short-term feeding tests. There are reliable analytical procedures available for virtually all of the known food-borne protein toxins and mycotoxins, which are used routinely in determining the safety of new production strains.

As of June 1999, members of the Enzyme Technical Association (ETA) reported conducting 102 bacterial mutagenesis tests (Ames *et al.*, 1975; OECD, 1984; EEC, 1992) and 63 chromosome aberration tests on enzyme preparations (Amacher *et al.*, 1980; OECD, 1984; EEC, 1992; Clive and Spector, 1975; Clive *et al.*, 1979, 1987). The chromosome aberration tests included *in vitro* cytogenicity tests in cultured mammalian cells (human peripheral lymphocytes, mouse lymphoma cells, or Chinese hamster ovary cells, for example) and *in vivo* tests in mice to detect damage to the chromosomes or the mitotic apparatus (OECD, 1984; EEC, 1992; Amacher *et al.*, 1980; Clive and Spector, 1975; Clive *et al.*, 1979, 1987). The enzyme preparations were from traditionally and genetically modified production organisms (i.e., 49 Ames tests and 27 chromosome aberration tests were conducted on enzyme preparations from genetically modified microorganisms). The production organisms were: *Actinoplanes missouriensis, Aspergillus melleus, A. niger, A. oryzae, Bacillus alcalophilus, B. amyloliquefaciens, B. licheniformis, B. naganoeneis, B. subtilis, Candida pseudotropicales, C. rugosa, Chaetomium erraticum, Disporotrichium dimorphosporum, Kluyveromces lactis, Leptographium procerum, Microbacterium imperial, Mucor javanicus, Penicillium camembertii, P. citrinum, P. decumbens, P. roqueforti, Pseudomonas alcaligenes, Rhizomucor miehei, Rhizopus niveus, R. oryzae (R. delemar), Streptomyces lividans, Talaromyces emersonii, Trichoderma reesei*, and *Verticidadiella procera*.

Seven of the Ames tests were false positive (6 on enzyme preparations from genetically modified microorganisms) and 6 of the chromosome aberration tests were false positive (2 on enzyme preparations from genetically modified microorganisms). The remainder (95 Ames and 57 chromosome aberration) were negative.

The false-positive results from the Ames tests were demonstrated to be due to the growth-enhancing effects of histidine in the enzyme preparations (this was verified by repeating the assay using the treat and plate method and observing no evidence of mutagenesis).

The false-positive results from the chromosome aberration assays have several explanations, as follows:

—Clastogenicity was observed in human lymphocyte cells; additional *in vitro* studies with Chinese hamster ovary cells were clearly negative (two studies/one from genetically modified microorganism);

—*In vitro* positive results were not confirmed by *in vivo* cytogenetic tests (three studies); and

—Some enzymatic reactions result in the production of hydrogen peroxide, which is known to cause clastogenic aberrations (one study/one from genetically modified microorganism).

These findings underscore the conclusion that testing enzyme preparations from traditional and genetically modified microorganisms for genotoxicity is unnecessary for safety evaluation.

A FOOD ENZYME SAFETY EVALUATION STRATEGY FOR THE 21ST CENTURY

New enzymes with improved properties are now being derived in numerous ways, including genetic modification (protein engineering) (Arnold and Volkov, 1999; Atwell and Wells, 1999; Cleland and Craik, 1996; Ford, 1999; Shaw *et al.*, 1999), protein "breeding" (Minshull and Stemmer, 1999), chemical modification (DeSantis and Jones, 1999), and by isolation from newly explored environments (Adams *et al.*, 1995; Demain and Davies, 1999; Hunter-Cevera, 1998; Madigan and Marrs, 1997;

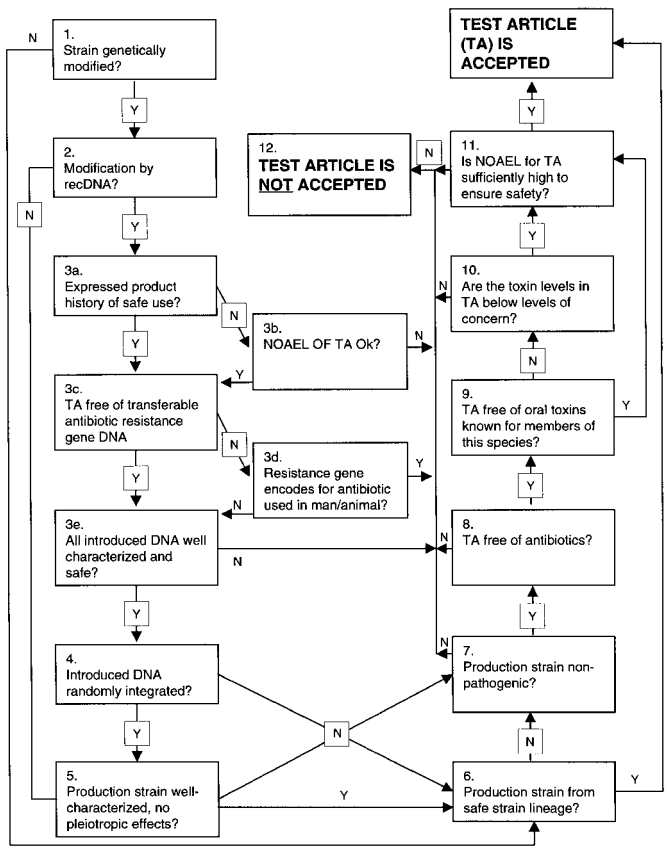


FIGURE 2

Marrs *et al.*, 1999; Pace, 1997). Accordingly, it is important to update previous enzyme safety evaluation mechanisms (Pariza and Foster, 1983; IFBC, 1990) to accommodate these advances in biology. The decision tree shown in Fig. 1 was developed to accomplish this goal. Figure 2 is a schematic representation.

The footnotes to the decision tree contain additional explanation and discussion. It is important to note that under the criteria of the decision tree, no new enzyme can enter the market without critical evaluation of its safety. It is also important to note that along with meeting the criteria in the decision tree, a safe food enzyme should also be produced under current Good Manufacturing Practices and meet or exceed the specifications for food enzymes described in the Food Chemicals Codex (FCC, 1996) and/or FAO/JECFA (JECFA, 1992).

CONCLUSION

Microbial enzymes used in food processing are typically sold as *enzyme preparations* that contain not only a desired enzyme activity but also other metabolites of the production strain, as well as added materials such as preservatives and stabilizers. The added materials must be food grade and meet applicable regulatory standards. In this report we present guidelines that can be used to evaluate the safety of the metabolites of the production strain that are also present in the enzyme preparation, including of course, but not limited to, the desired enzyme activity itself. This discussion builds on previous reports (Pariza and Foster, 1983; IFBC, 1990) and includes consideration of new genetic modification technologies, for example, modifying the primary structure of enzymes to enhance specific properties that are commercially useful.

The safety of the production strain should remain as the primary consideration in evaluating enzyme safety. The primary issue in evaluating the safety of a production strain is its *toxigenic potential*, specifically the possible synthesis by the production strain of toxins that are active via the oral route. *Pathogenic potential* is not usually an area of concern for consumer safety because enzyme preparations rarely contain viable organisms. Pathogenicity is, however, important to worker safety.

Thoroughly characterized nonpathogenic, nontoxigenic microbial strains, particularly those with a history of safe use in food enzyme manufacture, are logical candidates for generating a *safe strain lineage*, through which improved strains may be derived via genetic modification either by using traditional/classical or rDNA strain improvement strategies (IFBC, 1990). The elements needed to establish a safe strain lineage include thoroughly characterizing the host organism, determining the safety of all new DNA that has been introduced into the host organism, and ensuring that the procedure(s) that have been used to modify the host organism are appropriate for food use.

Enzyme function may be changed by intentionally altering the amino acid sequence (e.g., protein engineering). Examples include changing the pH optimum, increasing thermal stability, reducing the requirement for cofactors such as metal ions, and stabilizing the enzyme against chemical oxidation. It may be asked if such modifications might also affect the safety of an otherwise safe enzyme. To address this question we consider what is known about the natural variation in enzyme structure and function, and conclude that it is unlikely that changes which improve upon desired enzyme function will result in the creation of a toxic protein. It is prudent to assess such very small theoretical risks by conducting limited toxicological tests on engineered enzymes.

Despite the questions raised by Pariza and Foster (1983) regarding the scientific rationale and need for testing new food enzyme preparations for mutagenic activity, the practice continues, driven largely by regulatory requirements in some locales. Accordingly it is worth noting that, to our knowledge, the requirement that new enzyme preparations be tested for *in vitro* genotoxicity has failed to reveal the presence of a single mutagen or clastogen that would not have been detected using the more comprehensive decision tree approach described by Pariza and Foster (1983) and IFBC (1990).

The centerpiece of this report is a decision tree mechanism that updates previous enzyme safety evaluation mechanisms (Pariza and Foster, 1983; IFBC, 1990) to accommodate advances in enzymology. Under the criteria of this decision tree, no new food enzyme can enter the market without critical evaluation of its safety.

REFERENCES

Adams, M. W. W., Perler, F. B., and Kelly, R. M. (1995). Extremozymes: Expanding the limits of biocatalysis. *Bio/Technology* **13,** 662–668.

Aktories, K., and Just, I. (Eds.) (2000). *Bacterial Protein Toxins.* Springer-Verlag, Berlin.

Alouf, J. E., and Freer, J. H. (Eds.) (1999). *The Comprehensive Sourcebook of Bacterial Toxins*, 2nd ed. Academic Press, London.

Amacher, D. E., Paillet, S. C., Turner, G. N., Ray, V. A., and Salsburg, D. S. (1980). Point mutations at the thymidine kinase locus in L5178Y mouse lymphoma cells. II. Test validation and interpretation. *Mutat. Res.* **72,** 447–474.

Ames, B. N., McCann, J., and Yamasaki, E. (1975). Methods for detecting carcinogens and mutagens with the *Salmonella*/mammalian-microsome mutagenicity test. *Mutat. Res.* **31,** 347–364.

Archer, D. B., and Peberdy, J. F. (1997). The molecular biology of secreted enzyme production by fungi. *Crit. Rev. Biotechnol.* **17,** 273–306.

Arnold, F. H., and Volkov, A. A. (1999). Directed evolution of biocatalysts. *Curr. Opin. Chem. Biol.* **3,** 54–59.

Association of Manufacturers of Fermentation Enzyme Products (AMFEP) (1998). *Consumer Allergy Risk from Enzyme Residues in Food.* AMFEP, Brussels, Belgium.

Atwell, S., and Wells, J. A. (1999). Selection for improved subtilases by phage display. *Proc. Natl. Acad. Sci. USA* **96,** 9497–9502.

Balows, A., and Duerden, B. I. (Eds.) (1998). *Topley & Wilson's Microbiology and Microbial Infections. Vol. 2. Systematic Bacteriology.* Arnold, London.

Barbesgaard, P., Heldt-Hanson, H. P., and Diderichsen, B. (1992). On the safety of *Aspergillus oryzae. Appl. Microbiol. Biotechnol.* **36,** 569–572.

Beutler, F. H., and Cerami, A. (1997). Cachectin: More than a tumor necrosis factor—Mechanisms of disease. *N. Engl. J. Med.* **316,** 379–385.

Bonekamp, F. J., and Oosterom, J. (1994). On the safety of *Kluyveromyces lactis*—A review. *Appl. Microbiol. Biotechnol.* **41,** 1–3.

Bott, R., Dauberman, J., Caldwell, R., Mitchison, C., Wilson, L., Schmidt, B., Simpson, C., Power, S., Lad, P., Sagar, I. H., Graycar, T., and Estell, D. (1992). Using structural comparison as a guide in protein engineering. In *Enzyme Engineering XI* (D. S. Clark and D. Estell, Eds.), pp. 10–19. New York Acad. Sci., New York.

Brenner, S., Lewitter, F., Patterson, M., and Handel, M. (Eds.) (1998). *Trends Guide to Informatics.* Elsevier, Cambridge, UK.

Chu, F. S. Mycotoxins and alimentary mycotoxicoses. In press.

Claridge, M. F., Dawah, H. A., and Wilson M. R. (Eds.) (1997). *Species: The Units of Biodiversity.* Chapman and Hall, London.

Cleland, J. L., and Craik C. S. (Eds.) (1996). *Protein Engineering: Principles and Practice.* Wiley–Liss, New York.

Clive, D., and Spector, J. F. S. (1975). Laboratory procedures for assessing specific locus mutations at the TK locus in cultured L5178Y mouse lymphoma cells. *Mutat. Res.* **31,** 17–29.

Clive, D., Johnson, K. O., Spector, J. F. S., Batson, A. G., and Brown, M. M. M. (1979). Validation and characterization of the L5178Y/TK± mouse lymphoma mutagen assay. *Mutat. Res.* **59,** 61–108.

Clive, D., Caspery, W., Kirby, P. E., Krehl, R., Moore, M., Mayo, J., and Oberly, T. J. (1987). Guide for performing the mouse lymphoma assay for mammalian cell mutagenicity. *Mutat. Res.* **189,** 143–156.

Cohen, M. L. (1992). Epidemiology of drug resistance: Implications for a post-antimicrobial era. *Science* **257,** 1050–1055.

Conrad, B., Hoang, V., Polley, a., and Hoefmeister, J. (1995). Hybrid *Bacillus amyloliquefaciens* × *Bacillus licheniformis*-amylases: Construction, properties and sequence determinants. *Eur. J. Biochem.* **230,** 481–490.

Creighton, T. E. (1993). *Proteins: Structures and Molecular Properties*, pp. 105–138. Freeman, New York.

De Boer, A. S., and Diderichsen, B. (1991). On the safety of *Bacillus subtilis* and *B. amyloliquefaciens*: A review. *Appl. Microbiol. Biotechnol.* **36,** 569–572.

De Boer, A. S., Priest, F., and Diderichsen, B. (1994). On the industrial uses of *Bacillus licheniformis*: A review. *Appl. Microbiol. Biotechnol.* **40,** 595–598.

Demain, A. L., and Davies, J. E. (Eds.) (1999). *Manual of Industrial Microbiology and Biotechnology*, 2nd ed. ASM Press, Washington, DC.

DeSantis, G., and Jones, B. J. (1999). Chemical modification of enzymes for enhanced functionality. *Curr. Opin. Biotechnol.* **10,** 324–330.

Doolittle, R. F. (Ed.) (1996). *Computer Methods for Molecular Sequence Analysis, Methods In Enzymology*, Vol. 266. Academic Press, San Diego.

Doyle, M. P., Beuchat, L. R., and Montville, T. J. (Eds.) (1997). *Food Microbiology: Fundamentals and Frontiers.* ASM Press, Washington, DC.

EEC (1992). *EEC 1992 Guidelines—Annex V of the EEC Directive 67/548/EEC, Part B, Method for the Determination of Toxicity: B14 Bacterial Gene Mutation, Ames.*

ETA (1999). The enzyme preparations were from traditionally and genetically modified production organisms. The production organisms were *Actinoplanes missouriensis, Aspergillus melleus, A. niger, A. oryzae, Bacillus alcalophilus, B. amyloliquefaciens, B. licheniformis, B. naganoeneis, B. subtilis, Candida pseudotropicales, C. rugosa, Chaetomium erraticum, Disporotrichium dimorphosporum, Kluyveromyces lactis, Leptographium procerum, Microbacterium imperial, Mucor javanicus, Penicillium camembertii, P. citrinum, P. decumbens, P. roqueforti, Pseudomonas alcaligenes, Rhizomucor miehei, Rhizopus niveus, R. oryzae (R. delemar), Streptomyces lividans, Talaromyces emersonii, Trichoderma reesei*, and *Verticidadiella procera*.

Falkow, S. (1997) What is a pathogen? *ASM News* **7,** 359–365.

FDA (1993). 58 *Fed. Regis.* 27197, 27199 May 7, 1993.

FAO (1981). *FAO Food and Nutrition Paper: 25th Session of the Joint FAO/WHO Expert Committee on Food Additives*, Appendix A, pp. 317–318, FAO/WHO, Geneva, Switzerland.

Fink, G. R. (1999). What's in a name? *Cell* **96,** 468–469.

Fischetti, V. A. (Ed.). (2000). *Gram-Positive Pathogens.* ASM Press, Washington, DC.

Ford, C. (1999). Improving operating performance of glucoamylase by mutagenesis. *Curr. Opin. Biotechnol.* **10,** 353–357.

Geiser, D. M., Pitt, J. I., and Taylor, J. W. (1998). Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus. Proc. Natl. Acad. Sci. USA* **95,** 388–393.

Henrissat, B., and Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7,** 637–644.

Hunter-Cevera, J. C. (1998). The value of microbial diversity. *Curr. Opin. Microbiol.* **1,** 278–285.

International Food Biotechnology Council (IFBC) (1996). *Allergenicity of Foods Produced by Genetic Modification* (D. D. Metcalfe, R. L. Fuchs, R. Townsend, H. A. Sampson, S. L. Taylor, and J. R. Fordham, Eds.), *Critical Reviews in Food Science and Nutrition*, Vol. 36, Supplement. IFBC, Washington, DC.

International Food Biotechnology Council (IFBC). (1990). Safety evaluation of foods and food ingredients derived from microorganisms in biotechnologies and food: Assuring the Safety of Foods produced by genetic modification. *Regul. Toxicol. Pharmacol.* **12,** S1–S196.

International Life Sciences Institute (ILSI) (1997). *An Evaluation of the Budget Method for Screening Food Additive Intake.* ILSI, Washington, DC. [Also see http://www.ilsi.org/pubs/pilsieurope.html]

International Union of Bichemistry (IUB) (1992). *Enzyme Nomenclature: Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes.* Academic Press, San Diego. [Also see http://prowl.rockefeller.edu/enzymes/enzymes.htm]

Jancek, S., Leveque, E., Belarbi, A., and Haye, B. (1999). Close evolutionary relatedness of a-amylases from archaea and plants. *J. Mol.* Evol. **48,** 421–426.

Joint FAO/WHO Expert Committee on Food Additives (JECFA) (1992). *Compendium of Food Additive Specifications*, Vol. 1, *General Specifications for Enzyme Preparations Used in Food* Processing, pp. 1–3.

Kessler, D. A., Taylor, M. R., Maryanski, J. H., Flamm, E. L., and Kahl, L. S. (1992). The safety of foods developed by biotechnology. *Science* **256,** 1747–1749, 1832.

Klaassen, C. D. (1995). *Casarett & Doull's Toxicology: The Basic Science of Poisons*, fifth edition. McGraw–Hill, New York.

Kuchner, O., and Arnold, F. H. (1997). Directed evolution of enzyme catalysts. *Trends Biotechnol.* **15,** 523–530.

Kuhls, K., Lieckfeldt, E., Sammuels, G. J., Kovacs, W., Meyer, W., Petrini, O., Gams, W., Borner, T., and Kubicek, C. P. (1996). Molecular evidence that the asexual industrial fungus *Trichoderma*

reesei is a clonal derivative of the ascomycete *Hypocrea jecorina*. *Proc. Natl. Acad. Sci. USA* **93,** 7755–7760.

Kurtzman, C. P., and Fell, W. J. (Eds.) (1999). *The Yeasts: A Taxonomic Study*, 4th ed. Elsevier, The Netherlands.

Lehman, A. J., and Fitzhugh, O. G. (January 1954). 100-fold margin of safety. *Q. Bull. Assoc. Food Drug Off.* **18,** 33–35.

Madigan, M. T., and Marrs, B. L. (1997). Extremophiles. *Sci. Am.* **276,** 82–87.

Marrs, B., Delagrave, S., and Murphy, D. (1999). Novel approaches for discovering industrial enzymes. *Curr. Opin. Microbiol.* **2,** 241–245.

Mims, C. A. (1991). *The Pathogenesis of Infectious Disease*, 3rd ed. Academic Press, London.

Minshull, J., and Stemmer, W. P. C. (1999). Protein evolution by molecular breeding. *Curr. Opin. Chem. Biol.* **3,** 284–290.

Nevalainen, H., Suominen, P., and Taimisto, K. (1994). On the safety of *Trichoderma reesei*. *J. Biotechnol.* **37,** 193–200.

OECD (1984) *Guidelines for Testing of Chemicals, Section IV, Health Effects:* No. 471, Genetic Toxicology—*Salmonella typhimurium* Reverse Mutation Assay; No. 472, Genetic Toxicology—*Escherichia coli*, Reverse Mutation Assay; No. 473, Genetic Toxicology—*In Vitro* Mammalian Cytogenetic Test; No. 474, Genetic Toxicology—Micronucleus Test; and No. 476, Genetic Toxicology—*In Vitro* Mammalian Cell Gene Mutation Tests. OECD, Paris, France.

Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. Science **276,** 734–735.

Pariza, M. W., and Foster, E. M. (1983). Determining the safety of enzymes used in food processing. *J. Food Protect.* **46,** 453–463.

Patterson, D. S., and Roberts, B. A. (1979). Mycotoxins in animal feedstuffs: Sensitive thin layer chromatographic detection of a flatoxin, ochratoxin A, sterigmatocystin, zearalenone, and T-2 toxin. *J. Assoc. Off. Anal. Chem.* **62,** 1265–1267.

Pitt, J. I., and Hocking, A. D. (1997). *Fungi and Food Spoilage*, 2nd ed. Blackie Academic and Professional, London.

Rappuoli, C., and Montecucco, C. (Eds.) (1997). *Guidebook to Protein Toxins and Their Use in Cell Biology*. Oxford Univ. Press, Oxford.

Schlaes, D. M., Gerding, D. N., John, J. F., Jr., Craig, W. A., Bornstein, D. L., Duncan R. A., *et al.* (1997). Society for Healthcare Epidemiology of America and Infectious Disease Society of American Joint Committee on Prevention of Antimicrobial Resistance: Guidelines for prevention of antimicrobial resistance in hospitals. *Clin. Infect. Dis.* **25,** 584–599.

Scientific Committee for Food (1992). *Guidelines for the Presentation of Data on Food Enzymes*, 27th series, Ref. No. EUR14181 EN, pp. 13–22.

Shaw, A., Bott, R., and Day, A. G. (1999). Protein engineering of a-amylase for low pH performance. *Curr. Opin Biotechnol.* **10,** 349–352.

Siezen, R. J., and Leunissen, J. A. M. (1997). Subtilases: The superfamily of subtilisin-like serine proteases. *Protein Sci.* **6,** 501–523.

Strauss, E. J., and Falkow, S. (1997). Microbial pathogenesis: Genomics and beyond. *Science* **276,** 707–712.

Todd, A. E., Orengo, C. A., and Thornton, J. M. (1999). Evolution of protein function, from a structural perspective. *Curr. Opin. Chem. Biol.* **3,** 548–556.

Ward, D. M. (1998). A natural species concept for prokaryotes. *Curr. Opin. Microbiol.* **1,** 271–277.

Wolf, K. (Ed.) (1995). *Nonconventional Yeasts in Biotechnology: A Handbook*. Springer-Verlag, Berlin.

INDUSTRIAL ENZYMOLOGY

The Application of Enzymes in Industry

TONY GODFREY
and
JON REICHELT

M
The Nature Press

Published in the United Kingdom by
MACMILLAN PUBLISHERS LTD (Journals Division), 1983
Distributed by Globe Book Services Ltd
Canada Road, Byfleet, Surrey, KT14 7JL, England

ISBN 0 333 32354 8

Published in the USA and Canada by
THE NATURE PRESS, 1983
15 East 26th Street, New York, NY 10010

Library of Congress Cataloging in Publication Data
Main entry under title:
Industrial enzymology.
Includes index.
1. Enzymes – Industrial applications. I. Godfrey, Tony, 1939– . II. Reichelt, Jon, 1949–
TP248.E51516 1983 660'.63 82–14461
ISBN 0–943818–00–1

Printed in Great Britain

BREWING

T. Godfrey

1. Introduction

For many centuries the production of traditional fermented beverages by the extraction of cereals has been a mainstay of populations and governments. The technical skills of the brewer have never been in doubt, and the developments in the processes have been largely those of the engineering and packaging specialists. The traditional enzyme source for the conversions has been a variety of malted grains, but this chapter is not intended to provide an account of malting technology. If it is accepted that malted grain represents the mobilization of seed resources, via the synthesis and consequent action of hydrolytic enzymes resulting upon germination, then it can be readily appreciated that a number of enzymes will be present that have at least some contribution to make to the brewer's intentions. The germination process is arrested by heat and water removal, and provides a partly modified substrate together with dormant enzymes that can be activated, by grinding and mashing with warm water, to produce a typical brewer's mash. The main features of the conversion achieved with malted cereals can be described through the two main process methods of brewing.

2. Dominant malt enzymes

There are four enzymes in malts that are of primary benefit to the brewer; these are carboxypeptidases, α- and β-amylases, and β-glucanases. The upper limits of their thermal activity are set out in Table 4.5.1. It is clear that the next most important feature of malt brewing is the creation of conditions that utilize the enzymes to the maximum.

TABLE 4.5.1
Upper limits of the thermal activity of malt enzymes

Malt enzyme	*Thermal limit* (°C)*
α-amylase	68
β-amylase	64
β-glucanase	62
Carboxypeptidase	58

*Reduced to 20 per cent of original activity after 60 minutes.

This is largely achieved by attention to two parameters. First, enzymes are almost invariably stabilized towards adverse conditions of temperature and pH by being surrounded by an abundance of their substrate. The brewer therefore makes a mash of the highest concentration of ground grain that is compatible with the later separation of dissolved material from the mash; this is generally in the range of 28–35 per cent as dry solids. Second, a temperature for the mash is chosen that will either give a good compromise on the differing stabilities of the various enzymes, or a series of rising temperatures are selected which are optimal for the majority. These two quite different mashing methods are usually known as 'isothermal infusion' and 'programmed infusion' mashes.

3. Infusion mashing

Ground malt from barley, wheat, rye and sorghum will be converted to a fermentable material when mashed with hot water to give a mixed (strike) temperature of 63–65°C and a typical pH of 5.4. None of the enzymes released from the malt will be at its optimum temperature, but all will be protected to some extent by the starch and proteins released into the mash liquor. The key to the choice of temperature is that almost all of the starch should be fully gelatinized within 15–20 minutes, thus providing a substrate for the amylases. There is little evidence of proteolysis during mashing and it has been suggested that most of the observed protein solubilization has already occurred during malting. This can be confirmed by the thermolabile nature of the carboxypeptidases of malts, which will be almost completely inactive at the mash temperature.

Hydrolysis of β-glucans. This will similarly be limited by the temperature, and glucans are frequently present if the malt is poorly modified and low in β-glucanases.

Amylolysis. This is almost optimal in the infusion mash and both α- and β-amylases are active in converting the starch to a mixture of maltose and glucose, giving about 80 per cent fermentability of the original starch content. The remainder of the soluble starch products consist of higher dextrins and oligosaccharides that, whilst not being fermented, provide much of the characteristic body and mouthfeel of traditional beers. Figure 4.5.1 illustrates the action of these enzymes on starch components.

Malt α-amylase is an endo-enzyme with a random action that is limited to the outer linear sections of amylopectin. It is unable to degrade closer than a few glucose units from the α-1,6 branch points. Malt β-amylase, a sulphydryl exo-enzyme, also acts upon the linear amylose and linear portions of amylopectin to generate maltose units. It acts from the non-reducing ends of the chains.

The rate of these reactions is strongly dependent on temperature and it is very slow when acting upon ungelatinized starch. The infusion temperature is thus selected with reference to the gelatinization range of the total package of cereals in the mash (the combined materials being termed the grist). The ultimate fermentation character of the wort liquor will depend on the way these various factors of gelatinization and enzyme activity are combined.

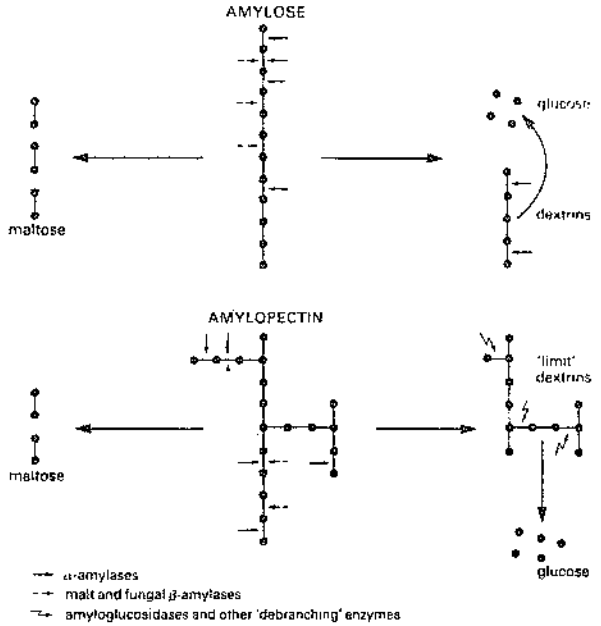


Figure 4.5.1 Schematic representation of the hydrolysis of starch components.

4. Programmed infused mashing

Basically, this series of stepwise rises in temperature has been adopted to reduce the total process time for wort production, and to permit the adequate utilization of poorer quality malts. Typically, steps are created at 52, 65, 68 and 75°C for varying times (*see* Figure 4.5.2). Up to 25 per cent of unmalted barley or cereal starches may be included in grists for this method of mashing, although extended times at some steps may be necessary to achieve satisfactory conversions in the absence of additional enzymes (*see* page 243, also Figure 4.5.3). Reference to Tables 4.5.1 and 4.5.2 shows that enzyme action, heat inactivation and gelatinization points will not occur in an ideal sequence, but a compromise is achieved that yields a satisfactory extract.

At the initial stage (52°C) a small amount of further proteolysis occurs, resulting in an increase in the soluble nitrogen levels to

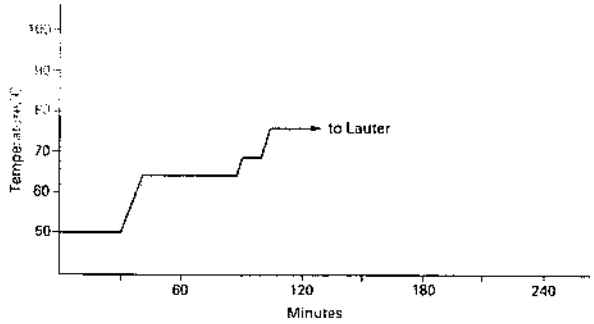


Figure 4.5.2 Mashing diagram: all malt infusion programme.

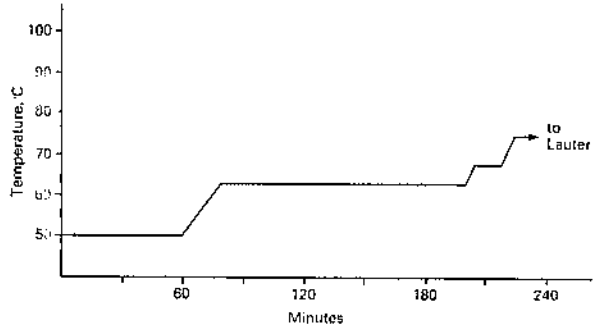


Figure 4.5.3 Mashing diagram: 30 per cent adjunct: 70 per cent malt infusion programme.

match the figures found when mashing with 100 per cent good quality malt. Glucanase attack upon the various gums of the grain begins here and continues throughout the heating up to stage 2. By the time mash reaches 65 °C, gelatinization of starch has commenced and the amylases are active. This process is completed during the rise and hold at 68 °C, and accelerates rapidly before being stopped by the rise to 75 °C.

TABLE 4.5.2
Gelatinization temperatures of various starches

Starch type	*Gelatinization temperature* (°C)
Maize (high amylose)	68–105
Maize	63–74
Rice	68–75
Sorghum	68–75
Barley malt	63–66
Rye	58–70
Wheat	55–65
Barley	53–58
Potato	53–60

5. Decoction mashing

This method of achieving conversion is characteristic of the production of lager beers, with a typical high percentage of unmalted cereals, and has also been used to facilitate the inclusion of low grade malts.

In addition to the basic infusion mashing procedures, a separate high temperature stage is included to facilitate the liquefaction of the starch adjuncts (*see* Figures 4.5.4 and 4.5.5). Traditionally, a portion of the malt is mixed with the adjunct in a separate cooker and then heated to boiling before being mixed back with the main mash. This hot material raises the temperature of the mash, which is then converted to fermentable products as already described. In some operations this process is repeated by removing a portion of the mash, adding more malt and boiling again before mixing back. In the USA and Canada it is common to use a 'double mash' system (*see* Figure 4.5.6).

The grist consists of a relatively small amount of malt, often as little as 35 per cent. The starchy adjuncts consisting of maize grits or rice are mashed in the 'cereal cooker' with an initial hold at 35 °C for up to 60 minutes, to activate the small amount of malt enzymes.

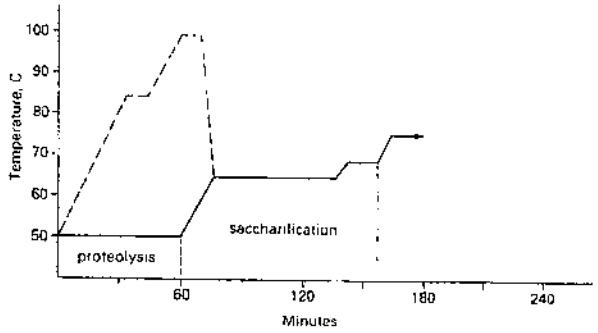


Figure 4.5.4 Mashing diagram: single decoction adjunct cooking with malt mash.

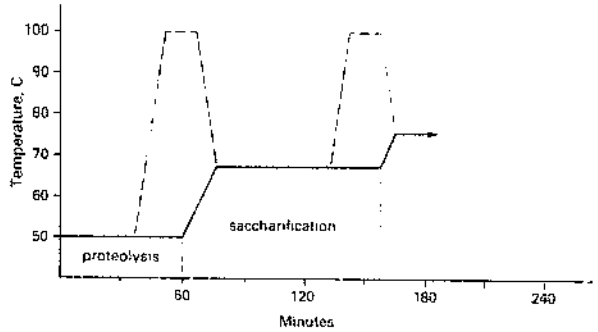


Figure 4.5.5 Mashing diagram: double decoction adjunct cooking with malt mash.

This adjunct mashing is designated the 'first mash' and is heated to 70 °C and held for 20 minutes at this temperature for saccharification. It is then further heated to boiling and maintained there for 45 minutes. When this last heating stage has begun, the 'second mash', consisting entirely of malt, is initiated at 35 °C. Finally, the boiling first mash is combined with the malt mash to give a strike temperature of 65–68 °C. Saccharification now proceeds on all the released starch materials, and about two hours after the start of the process the mash is heated to 73–75 °C and pumped to the Lauter.

The introduction of microbial enzymes as a replacement for the malt that is lost in the decoctions has made a positive contribution to the economics of this method of brewing (*see* page 234).

6. Economic brewing with additional enzymes

The rising costs of malt production and the variations in malting quality of the barley, due to choice of variety and the weather during the growing season, have encouraged the brewer to include a number of additional enzymes in his process. Adjuncts provide a cheaper extract but lack the necessary enzymes for their conversion, and without additional enzymes their use rates are limited by the amount of malt needed to process them. These extra enzymes are available as standardized products that can give accurate control of the brewing process, provide flexibility for the brewer in the choice of grist components and reduce the overall cost of the materials for beer production.

There are many commercially available industrial enzymes for the brewing process, but they fall broadly into three categories: proteases, amylases and glucanases. Fungal, bacterial and plant sources are used for their production and many different aspects of the processes in the brewery are involved. A general guide to the

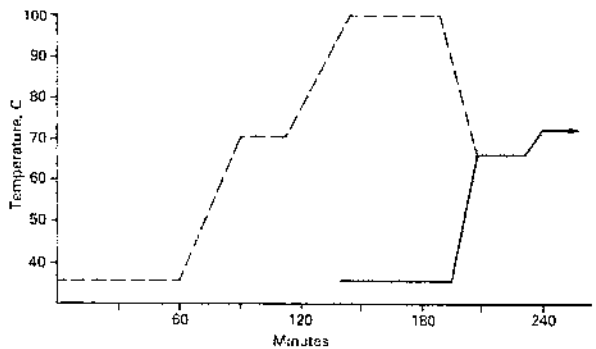


Figure 4.5.6 Mashing diagram: North American 'double mash' adjunct cooking with malt.

types of enzymes and their points of influence is given in Table 4.5.3. In order to follow the basis for the selection of enzymes to aid the brewing process, each stage in the natural sequence converting raw materials to finished beer will be considered in turn, according to the very simple chart set down in Fig. 4.5.7.

TABLE 4.5.3
Typical exogenous enzymes applied to brewing

Enzyme type	*Beneficial action*	*Point of application*
Bacterial α-amylases	Adjunct liquefaction	Decoction vessel (cereal cooker)
	Adjunct liquefaction Malt improvement	Mash vessel
	Set mashes	Mash vessel
	Starch positive worts	Lauter or mash filter
Fungal α-amylases (maltogenic action)	Improved fermentability	Fermentation
	Low calorie and 'diet'	Fermentation
	Set mashes	Mash vessel
	Starch positive worts	Lauter or mash filter
Fungal amyloglucosidases	Low calorie and 'diet'	Fermentation
	Maximum fermentability	Fermentation
	Priming replacement	2° fermentation or post-pasteurization
Bacterial debranching enzymes	Maximum fermentability	Fermentation
Bacterial glucanases	Increased extract	Mash vessel
	Improved wort separation	Mash vessel
	Improved filtration	Mash vessel/fermentation/ conditioning tank
Fungal glucanases (including cellulases)	Improved extraction	Mash vessel
	Improved wort separation	Mash vessel
	Improved filtration	Mash vessel/fermentation/ conditioning tank
	Increased adjunct (especially sorghum)	Mash/decoction vessel
	Haze prevention	Mash vessel
	Haze removal	Fermentation/ conditioning tank
Bacterial neutral protease	Increased adjunct	Mash vessel
	Nitrogen regulation	Mash vessel/fermentation
Plant proteases (papain)	Chillproofing against protein hazes	Conditioning tank
Fungal pentosanases	Prevention/removal of specific haze components	Mash vessel/fermentation/ conditioning tank
	Improved extract (especially wheat and sorghum)	Mash vessel

In the initial mashing stages, the selected grains and malts are steeped in warm water, either separately or together, and ultimately

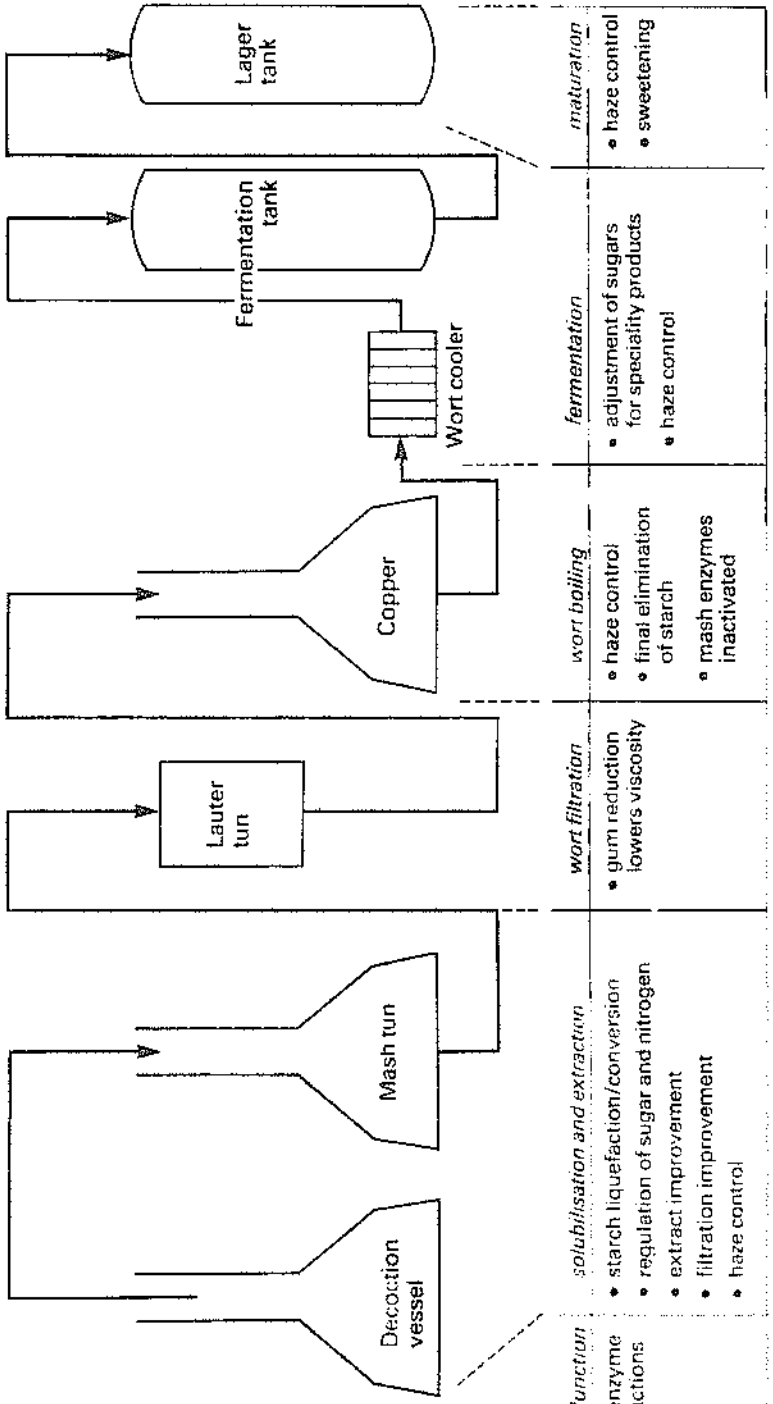


Figure 4.5.7 Operational sequences of the brewing process indicating the principal influences of enzyme action.

combined in a vessel designed to facilitate the separation of soluble from insoluble material. Modern methods for this separation include the traditional Lauter, mash filter systems and also centrifugation. The resulting fluid containing the extracted soluble material from the grains is termed the 'wort', and is destined to become the fermentation medium for the beer production.

The second main stage is the boiling of the wort, with or without hops or their extracts. The chemical changes characteristic of the flavour of hopped beers are active during boiling. In addition, this heat treatment is used to concentrate the wort to a level that will give sufficient fermentable sugars to yield the desired ethanol content for the final beer. A third function of the boiling stage is the sterilization of the wort to provide a clean medium for the yeast fermentation. This last effect includes the destruction, by thermal denaturation, of the malt and any added enzymes. This is to some extent a positive control step, ensuring that no further changes in the chemical composition of the wort will occur unless directly as a result of yeast fermentation or subsequently added enzymes. For many breweries utilizing a number of additional enzymes in the mashing stages, this is also confirmation that the enzymes are no longer active and not therefore likely to be subject to complicated labelling regulations. However, this does not remove responsibility for the selection of 'approved' additives at all stages in the process. These factors have been of great significance in the increasing drive to restrict adjustments to the brewing process to the mash vessels. There are certain specific cases, however, such as the production of fully fermented 'diet' products or the treatment of hazes appearing at later stages, where it can be beneficial to make further enzyme additions so as to reintroduce specific activities (*see* pp. 252, 255).

The cooled wort is then fermented with selected yeasts under controlled conditions to produce the required alcohol levels, and some beers then undergo a period of maturation, usually after the removal of the bulk of the yeast, to complete the development of specific product character.

The following sections will illustrate the various enzymes applied to each main step in this brewing process. There are many instances where more than one stage can be the point for a particular target change to be accomplished, and in many cases the same activity can be used at different stages for slightly differing targets.

7. Malt mashing improvement

Although many different factors influence the performance of malted cereals, there are three very general factors for the brewer to consider: (*i*) the type of barley (or other malting cereal), and the

manner of its cultivation, together with the vagaries of seasonal weather; (*ii*) the method of malting and any additional chemical treatments at the maltings; (*iii*) the mashing system selected by the brewer.

For the first two of these factors, economic considerations coupled with market forces may often combine to produce a grain of much reduced quality. The professional maltster can improve on the quality by making various adjustments to the malting process, perhaps by altering the steeping times, or adding plant hormones to stimulate the germination, and monitoring the influence of his kilning programme on the quality of the extract obtained.

In this last factor, the brewer chooses the malt to suit his method or evaluates the available malt in terms of its economy together with the use of exogenous enzymes. These enzymes are added to the mash to supplement malt enzymes of insufficient potency, and in many cases to provide additional activities not inherently present in even the best quality malts.

Undermodified malts. These are characterized by their lower amounts of total extractable soluble material and lower glucanase and amylase activities than their high quality counterparts. Worts produced unaided from them often have a higher viscosity, with slower run-off and often smaller wort volumes, as a result of the larger amounts of unhydrolysed glucans present, together with lower soluble nitrogen components and poorer fermentability. These various shortcomings can be tackled by the selection of appropriate mashing enzymes. Microscopic examination of the mash often shows that considerable amounts of starch remain trapped in the fibres of the cell debris. These starch particles can be released by the addition of fungal cellulases, which hydrolyse β-1,4 bonds of both cellulose and hemicellulose fibres of cereals. Extract improvements of up to 15 per cent can be obtained by the addition of as little as 0.2 kilogrammes *Celluclast*® 2.0L (Novo) per tonne of malt. Furthermore, the glucanase action of the cellulases generally results in improved filtration of the wort. However, the depressed levels of malt glucanases so characteristic of poor malts are aggravated by the usually elevated levels of glucans in these malts. The most economic improvement in this imbalance is usually achieved by the addition of combinations of bacterial and fungal glucanases other than the cellulases. Both types hydrolyse β-1,3 and -1,4 bonds of glucans but with differing endproducts. The bacterial enzymes generate small oligosaccharides of three to seven glucose units, but cause a dramatic reduction in wort viscosity, while the fungal types yield glucose with a wide range of residual oligosaccharides and useful reduction in viscosity. Practical experience shows that the use

of one of each type of these enzymes is far more effective than an increased dose rate of either alone. There is little to choose between the various bacterial glucanases, but quite marked differences are seen in the action of the same activity dose of different fungal glucanases. An appropriate selection may easily be made on the basis of small-scale laboratory mashes with the chosen grist materials. Generally, it is found that the bacterial glucanases are used at twice the rate of the fungal glucanases, assuming they are all on the same activity scale.

The fermentability of the wort is determined after its separation from the grains. The value is nevertheless determined by the degree of saccharification of the mash, and this in turn depends upon the activity of the glucogenic and maltogenic enzymes present. Whilst these are generally not found to be limiting in even very undermodified malts, the restricted release of starches from the cell matrix, described above, can slow down the overall development of fermentable material to a point where the mashing cycle is disrupted. The dextrinizing activity of the mash is thus frequently supplemented with bacterial α-amylase. Rapid hydrolysis of α-1,4 bonds in amylose and amylopectin by the addition, for example, of up to 1 kilogramme *BAN* 120L (Novo) per tonne malt will provide an improved substrate for the malt enzymes and re-establish the fermentability of the extract.

To produce worts equivalent to high quality malts, the amount and distribution of nitrogen compounds may need adjustment. The ratio of peptide to α-amino nitrogen should be maintained when increased protein hydrolysis is introduced. Although there are many proteases apparently suited to this application, many of them are susceptible to inhibitors produced by raw barley and often present in mashes of undermodified malts. The low levels of protein in the mash, coupled with generally adverse temperatures and often slightly sub-optimal pH values, tend to generate unstable conditions for even those proteases that can function otherwise adequately in the mash. In programmed mashes the stand at 52–55 °C helps to give a greater dose–time response for proteolytic action. In isothermal mashing, it should be assumed that the proteases will only remain at above 50 per cent of their added activity for about 30 minutes. The enzyme *Neutrase*® 0.5L (Novo), used at 0.3–1.5 kilogrammes per tonne malt, can raise the soluble nitrogen levels very effectively and maintain the required ratios, as demonstrated in Figure 4.5.8. It is not normal to raise the nitrogen levels by more than about 20 per cent over the untreated values to correct for undermodified malts. Some typical malt improvement enzymes and doses are given in Table 4.5.4.

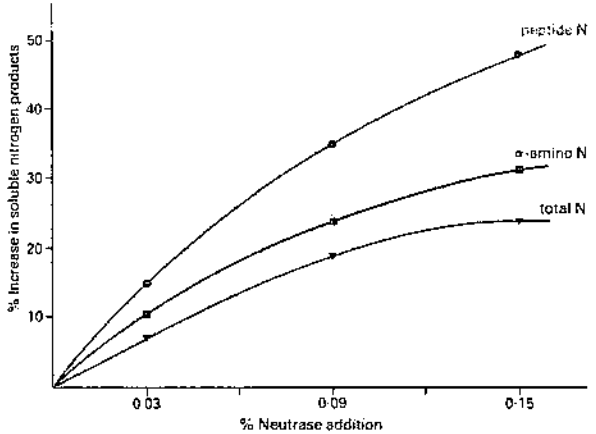


Figure 4.5.8 Increase of soluble nitrogen compounds by the application of *Neutrase*® 0.5L to a malt mash.

TABLE 4.5.4
Enzymes for malt mash improvement

Enzyme type	*Example product (Novo)*	*Use rate (kg per tonne malt)*
Bacterial α-amylase	*Bacterial Amylase Novo* 120L	0.5–1.0
Bacterial β-glucanase	*Cereflo*® 200L	0.5–1.0
Fungal β-glucanase	*Finizym*® 200L	0.2–0.5
Fungal cellulase	*Celluclast*® 2.0L	0.2–0.4
Bacterial proteinase (neutral)	*Neutrase*® 0.5L	0.3–1.5

8. Adjunct processing

The considerable economic advantage of the extract obtained by adding unmalted raw materials has encouraged their inclusion in the mash bills of most breweries. In addition to the lower cost of the extract, they contribute to the regulation of several other features of beer production. These include provision of fermentable compounds without adding tannins or proteins, reduction in beer colour whilst maintaining fermentable gravity, and improvement in beer stability and foam character.

The choice of adjunct depends as much on the country, availability and price as upon the particular brewing change intended, but would generally be made from the following: maize grits, wheat flour, rice, sorghum, starches from these and other sources, and of course, sugar syrups from starch conversions and sucrose. As these last two groups are fully processed products, there is no need to consider them here except to refer the reader to Chapter 4.15, 'Starch', for a discussion of the production of enzyme-converted

starch sugars.

Some commonly used brewing adjuncts are given in Table 4.5.5, with additional indications of their protein content and benefit from enzymic processing. Although the first four adjuncts have a significant protein content, this is not generally extracted at the brewery (but *see* page 231 for the adjustment of protein). Adjuncts 5, 6 and 7 are substantially free of proteins so that the first 7 can all be used to provide fermentable products without adding either proteins or tannins to the wort. The last two adjuncts are used to replace malt but also to retain the wort composition. Barley will be discussed more fully on page 240.

Regulation of the levels of proteins and tannins in beer is very important in preventing the formation of hazes, as well as for the effects of chilling of the finished beers. Elevated levels of a group of substances described as glycoproteins are claimed to enhance the foaming character of beer and it is often beneficial to add wheat flour or barley, which readily contribute such substances.

Enzymic processing of adjuncts. In all cases, starchy adjuncts require gelatinization, liquefaction, dextrinization and saccharification to convert them to fermentable products. Table 4.5.2 illustrates the wide variation in gelatinization temperatures needed for the different starch sources, but it should be noted here that 'flaked' cereals are processed prior to sale to breweries. The treatments include heating to temperatures that are generally in excess of the gelatinization minimum, which permits them to be processed directly in the mash tun together with the malt grist. However, if the rate of adjunct use is to exceed the conversion power of the malt content, then even flaked cereals may require either separate cooking or the addition of the same range and doses of enzymes as

TABLE 4.5.5
Protein content and enzyme processing benefit of adjuncts

Adjunct	*Protein content*	*Process benefit when enzymes used*
1 Maize	+	+
2 Wheat flour	+	+
3 Rice	+	+
4 Sorghum	+	+
5 Starch (various)	±	+
6 Starch syrups	–	–
7 Sucrose	–	–
8 Malt extract	+	–
9 Barley	+	+

required for malt improvement (*see* page 229).

Adjunct liquefaction and dextrinization. Separate cooking in the decoction (cereals) cooker is used to raise the adjunct to gelatinization temperature without heating the whole malt mash to too high a temperature for the survival of the malt enzymes.

The traditional source of liquefying amylases is malt, but the proportion used for this stage is lost by the boiling. To avoid this loss and also to improve the overall liquefaction rate, it is common practice to use bacterial amylases for the cooking. The conventional use of amylases derived from *Bacillus subtilis* has provided economic improvements over malt for three decades, but recently the introduction of the more heat-stable amylases from *Bacillus licheniformis* has produced still further improvements.

Table 4.5.6 shows the essential features of these two types of amylase and the ways in which their performance differs. Improved yields from adjuncts with small starch granules are the result of more complete gelatinization at the higher temperatures made possible by these amylases.

Conventional amylases. These require a hold of 20–30 minutes at around 70–75°C in the heat-up cycle of the cooker. Typically, doses of two to three kilogrammes of, for example, *Bacterial Amylase Novo* 120L per tonne starch content would be added at the charging of the cooker. These amylases also show a marked requirement for the presence of free calcium for their stable action. Calcium hydroxide is frequently added to the mash before cooking to raise the pH to around 6.5 and raise the calcium levels to the 150–200 parts per million required. The boiling stage of the cooking also inactivates the amylase so that, although full gelatinization may be achieved, there is always the possibility of incomplete dextrinization of the starch. This is completed in the subsequent malt mashing with the cooked adjunct, but can be a limitation on the upper levels of adjunct that can be used overall.

TABLE 4.5.6

Comparative operating characteristics of conventional and thermostable bacterial α-amylases for adjunct liquefaction

Amylase	*Conventional*	*Thermostable*
Source	*B. subtilis*	*B. licheniformis*
Optimum operating temperature	80–85°C	95–105°C
Optimum operating pH	6–6.5	6–6.5
Minimum calcium for stability	150 ppm	70 ppm
Upper starch concentration for practical thinning	30%	45%
Typical dose level	2–3 kg per tonne	1–1.5 kg per tonne

Thermostable amylases. These may be used at much smaller dose levels, for example 1.0–1.5 kilogrammes *Termamyl®* 60L (Novo) per tonne starch content, and without the addition of any calcium salts except where the starch and water supply are both very low in calcium, so that less than 70 parts per million would be present in the mash. Such a combination of circumstances is very uncommon.

To reduce cooking times, it is also possible to mash the adjunct directly with hot water at, say, 90°C. In all cases the cooker may be rapidly heated directly to the boil and maintained for 30 minutes to complete the liquefaction. At the end of the cooking cycle the mash viscosity is generally down to a pumpable 400 centipoise and the mash is being rapidly dextrinized under the action of the amylase that will have survived to a large extent. Figures 4.5.9, 4.5.10 and 4.5.11 show examples of the mashing programmes and viscosity characteristics for the different liquefying enzymes.

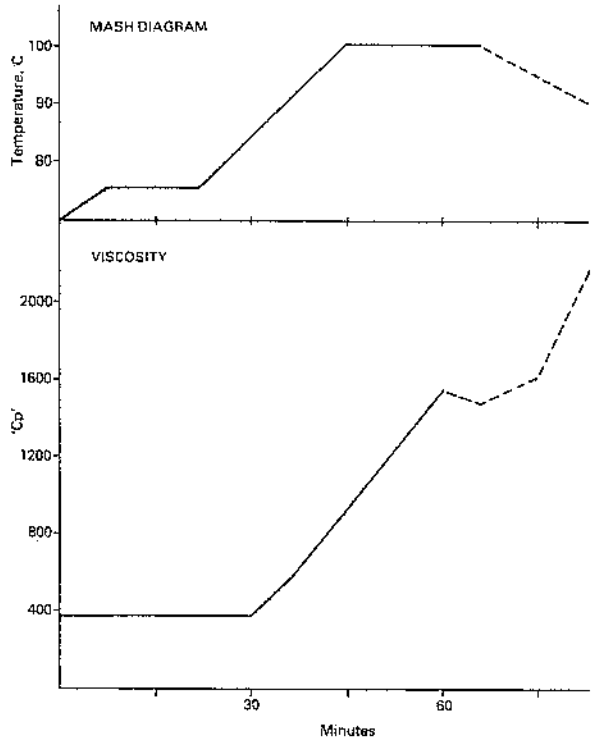


Figure 4.5.9 Liquefaction of maize grits with malt.

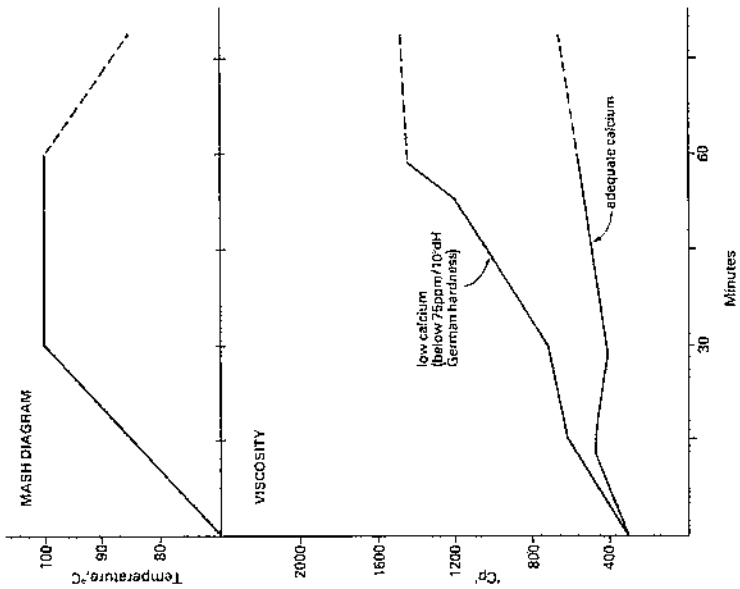


Figure 4.5.11 Liquefaction of maize grits with thermostable amylases.

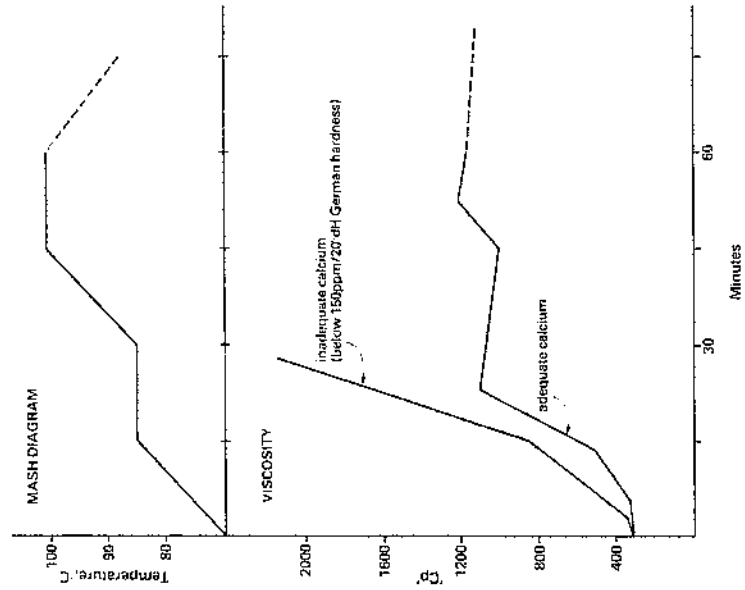


Figure 4.5.10 Liquefaction of maize grits with conventional amylases.

Adjunct mash concentrations. The thermostability scale for these enzymes approximates to their liquefying performances, which can be translated into increasing cooker capacity potential. The typical malt enzyme adjunct cook will be operated at 25 per cent solids in the cooker; conventional bacterial amylases (with adequate calcium) will act at up to 32 per cent solids, while the thermostable amylases will run up to 40 per cent solids, thus making it practical to increase the capacity of the cooker by up to 60 per cent.

Continuous cooking of adjuncts. A further property of the thermostable amylases is the survival of almost 100 per cent of their activity when they are briefly exposed to high temperatures in the presence of stabilizing levels of starch substrates. Whilst primarily a technique for the starch syrup industry (*see* Chapter 4.15), the application of the continuous cooking of starch adjuncts in brewing has been increasing.

The process can utilize a simple stirred tank cooker, but is at maximum efficiency when a pressurized jet cooker (such as those supplied by the Hydrothermal Corporation of Milwaukee, USA) is used. The jet system consists of separate feeds of starch slurry and clean steam to a venturi device that creates a combination of instant heating and shearing turbulence. The slurry is usually maintained at elevated pressure for 5–10 minutes in a holding tube before being flashed down to atmospheric pressure into a further holding tank for 30–45 minutes to ensure dextrinization of the starch. Energy savings on the cooker system are reported to be up to 25 per cent compared with batch cooking. The main criterion for the adjunct for this method of cooking is that it be finely ground as a flour so that it can pass freely through the jet system. Any cereal flour may be jet cooked with advantage using equipment based on the system shown in Figure 4.5.12.

In this system, a slurry of the adjunct in water to the chosen solids concentration, typically 35–40 per cent, is fed to the jet from a stirred holding tank into which the enzyme is metered. Using *Termamyl*® 60L, the jet cooker is operated at a pressure sufficient

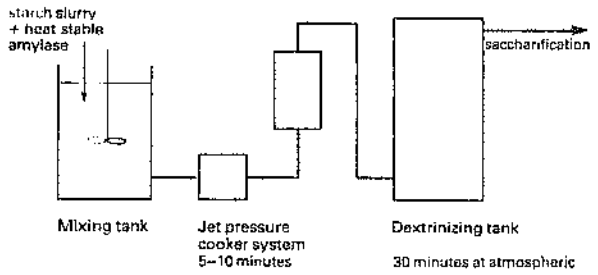


Figure 4.5.12 Schematic layout of a typical jet cooking conversion plant.

to generate a temperature of 105–110°C in the holding cell where the slurry is liquefied for 5–10 minutes. A dose of 1.5 kilogrammes *Termamyl®* 60L per tonne starch solids will result in a fully liquefied and iodine-negative dextrin slurry after a final hold of 30 minutes at atmospheric pressure and 95°C.

Special considerations for processing rice. In batch cooking of rice, as flour or as broken rice grits, gelatinization and liquefaction can be achieved at very much reduced enzyme levels. Typically, only 50 per cent of the levels needed for maize will produce excellent results; for example, *Bacterial Amylase Novo* 120L at 1–1.5 kilogrammes per tonne starch content, and *Termamyl®* 60L (Novo) at 0.5 kilogrammes per tonne starch content (*see* Figure 4.5.13).

Special considerations for processing sorghum and wheat. The presence of fibres and cellular debris, including in the case of wheat a considerable amount of pentosan polymers, can prevent the cooking systems from acting at optimal efficiency. A programmed batch cooking of these adjuncts, to include hydrolysis by cellulases prior to the high temperature starch liquefaction, has been found to enhance the processing characteristics of the subsequent combined adjunct and malt mashes. Increased levels of these adjuncts can be easily introduced by this process. Figure 4.5.14 illustrates the system for sorghum, but it is equally applicable to wheat. An initial dosing of the mash with, for example, *Termamyl®* 60L (Novo) at 1 kilogramme per tonne of adjunct starch, together with *Celluclast®* 2.0L (Novo), also at 1 kilogramme per tonne, is followed by a 20-minute hold at 60°C. During this period, the thermolabile cellulase is acting. The temperature of the cooker is then raised to boiling and held for 30 minutes for the starch to liquefy as in the previous examples. To transfer this system to continuous cooking simply utilizes the jet feed tank to give the cellulase treatment at as high a temperature as possible without inducing gelatinization of the selected adjunct (*see* Table 4.5.2).

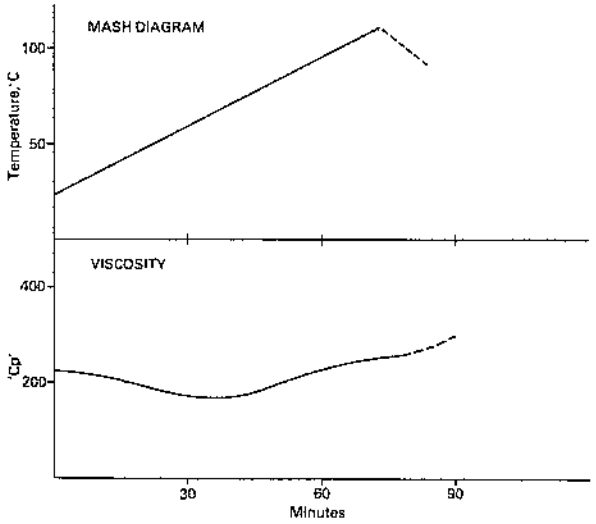


Figure 4.5.13 Liquefaction of cracked rice with *Termamyl®*.

Saccharification of dextrinized adjunct mashes. At modest levels of adjunct utilization, it is normal practice to combine the decoction mash with the malt mash and to achieve saccharification to fermentable compounds by the action of the malt enzymes. When higher adjunct levels are being used for the production of lighter coloured beers and 'diet' products, it may be practical to perform the adjunct saccharification separately.

The mash in the decoction vessel, or dextrinizing tank of the jet system, is cooled to a temperature suited to the saccharifying enzyme to be used, and the pH is adjusted downwards to the optimum for the chosen enzyme or enzymes, which are then added. The choice of enzymes depends on the spectrum of sugars one wishes to obtain, which range from very high levels of glucose, through a mixture of glucose and maltose to a low glucose–high maltose pattern. Fungal amyloglucosidases will degrade most of the dextrins resulting from the cooking stage down to glucose and also degrade many of the branched dextrins having α-1,6 branch bonds. Fungal β-amylases, on the other hand, have a much more limited degradation of these branched dextrins, but are highly specific for

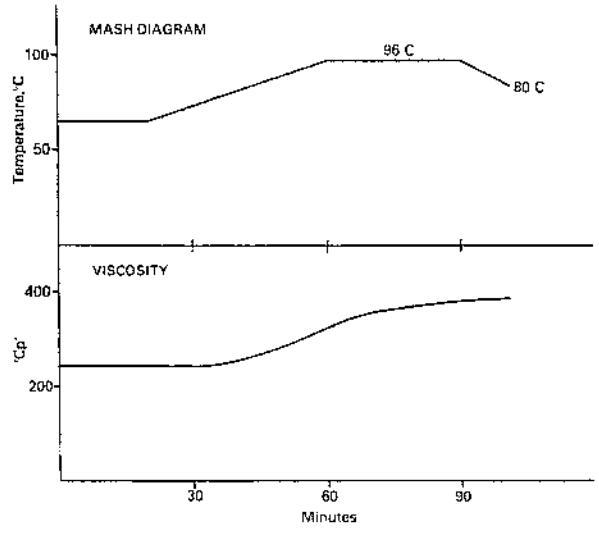


Figure 4.5.14 Liquefaction of sorghum starch with *Termamyl®* and *Celluclast®*.

the production of maltose as the predominant fermentable saccharide.

By selecting one of these enzymes, or a combination of the two, it is possible to generate a variety of sugar spectra with different characteristics of fermentation. Some of these possibilities are illustrated in Table 4.5.7, using the enzymes *Amyloglucosidase Novo* 150L and *Fungamyl*® 800L (Novo) at the levels indicated. These levels are far from rigid, and the dose/time/proportion of enzyme used should be determined by practical observation with the adjunct of choice.

After the desired saccharification stage has been reached, the adjunct is either filtered independently and passed to the copper, or added to the finished malt mash for separation of wort using the method preferred by the brewery.

Enzyme inactivation considerations. Of the enzymes mentioned in this section, the commercially available cellulases, conventional amylases and fungal β-amylases are all thermolabile and are inactivated at the cooker stage, or when the mash is heated before separation. The commercial amyloglucosidases so far investigated have a stability that requires heating to at least 82 °C for 30 seconds to inactivate them. Thus the enzymes used for malt improvement (*see* page 232) and those for adjunct processing will be inactivated during the next stage, that of wort boiling in the copper. Detailed studies of the deactivation of even the thermostable amylases such as *Termamyl*® have confirmed that they, too, are rapidly inactivated under copper boiling conditions. It is likely that the almost complete absence of substrate by this stage, coupled with reduced pH, enhances the deactivation of the enzyme by removing two of the critical stabilizing factors for these enzymes. After 30 minutes' boiling of wort, the activity of *Termamyl*® is reduced to below 20 per cent of its initial level. At the end of normal wort boiling, the activity has fallen to zero.

Nitrogen extraction at high adjunct levels. Despite the primary benefit of reduced nitrogen content by the use of adjuncts of virtually zero nitrogen contribution, it is usually necessary to increase the release of soluble nitrogen from the malt part of the mash when adjuncts are to exceed 30 per cent. It is of practical benefit, therefore, to use bacterial neutral proteinases in the malt mashing stage in a similar manner to that used for malt improvement as described earlier (*see* page 231). Table 4.5.8 summarizes some typical adjunct processing enzymes and their use rates.

9. Brewing with barley

Building on the earlier work of various brewing researchers, several

TABLE 4.5.7
Variations of sugar spectrum from adjunct saccharification with enzymes and subsequent mashing with malt (30% adjunct)

Amyloglucosidase (e.g. AMG 150L (Novo), % w/w of adjunct)	*Fungal β-amylase (e.g. Fungamyl® 800L, % w/w of adjunct)*	*Temperature (°C)*	*Time (hours)*	*% Sugar distribution*			
				Glucose	*Maltose*	*Maltotriose*	*Dextrins*
0.15	—	60	1.0	36	30	12	22
0.10	0.075	60	1.5	23	41	13	23
—	0.25	55	1.0	4	52	20	24
—	0.10	55	1.5	3	54	21	22

TABLE 4.5.8
Enzymes for adjunct processing

Enzyme type	*Example product (Novo)*	*Use rate per tonne malt*		*Point of addition and function*
		Maize	*Other*	
Conventional bacterial α-amylase	*Bacterial Amylase Novo* 120L	2–3 kg	1–1.5 kg	Cooker, liquefaction
Thermostable bacterial α-amylase	*Termamyl*® 60L	1–1.5 kg	0.5–1.0 kg	Cooker, liquefaction
Fungal cellulase	*Celluclast*® 2.0L	—	0.8–1.2 kg	Cooker, cellular breakdown
Fungal amyloglucosidase	*Amyloglucosidase* 150L		1.0–1.5 litres	Converter, saccharification
Fungal β-amylase	*Fungamyl*® 800L		0.75–2.5 kg	Converter, saccharification
Bacterial neutral proteinase	*Neutrase*® 0.5L		0.3–1.5 kg	Malt mash, nitrogen control

reports of enzyme-assisted barley brewing appeared in the early 1970s. The European Brewery Convention Congress at Estoril in 1971 provided a forum for much of the discussion and exchange of data that led, by 1973, to the accumulation of considerable practical evidence for the economic advantages of using raw barley as a major source of extract. Developmental studies and large-scale trials were reported frequently. By the time that the European harvest was to be so severely distressed by the drought of 1976, a considerable body of data existed to allow a confident brewer to embark on this route to economic brewing. In that and the subsequent year raw barley was to provide some 20 per cent of the extract in UK beers. In some subsequent years the overall proportion declined slowly to a steady level of approximately 15 per cent of brewer's extract, with a parallel gradual introduction of the technology into other European countries and recently into North America.

The most generally accepted method uses an infusion with an upward temperature gradient in agitated mash vessels. Interestingly, some of the main manufacturers of brewery equipment have designed plant especially suited to the conditions required for barley brewing and have installed them in many major breweries.

Barley forming up to 80 per cent of the grist can be used, but the most common levels are from 30 to 60 per cent, and some consideration to the milling is needed, as barley is much harder than malt. Various schemes have been adopted to make economic milling practicable, and these include the adjustment of malt mills to a more compromising tolerance and the insertion of harder steel rolls. The most successful method seems to be to arrange for a premilling steeping of the combined malt and barley grist followed by wet-milling.

Mashing conversion is required to produce a final wort that gives a finished beer with the character and quality of that from grists with a much larger malt content. Whilst the starch conversion chemistry of microbial enzymes is readily appreciated by the brewer familiar with adjunct processing, the adequate nutrition of the yeast and the elimination of filtration limiting and haze-forming potentials from the process are not so well known. The use of a range of enzymes in the barley mashing system can stabilize these various factors and provide substantial economic benefits for this form of extract. The processing of raw barley, with a more varied range of industrial enzymes than generally required for the other adjuncts, enables the brewer to produce a wort very closely related to that from an all-malt grist.

Enzymes for mashing barley. The addition of the microbial

equivalents for the malt proteinases, amylases and glucanases has proven a practical scheme since, in particular, barley has high levels of β-amylase and a protein content that is a good substrate for yielding soluble nitrogen precisely like that of malt. Barley starch also gelatinizes at mashing temperatures and therefore does not need to be cooked separately from the malt. A typical barley–malt mashing programme is illustrated in Figure 4.5.15. Proteolytic action is developed in the first part of the programme at 50°C. The enzymes that are entirely of the neutral and uninhibited type are essential here if the maximum benefit is to be obtained, since many of the mixed proteases have undeclared amounts of alkaline optimum activity that is inhibited by barley. An example of the preferred type is *Neutrase*® 0.5L from Novo, which is used at levels of 1.2–1.5 kilogrammes per tonne barley. At this level there will be sufficient for both the release of nitrogenous soluble compounds and the activation of the β-amylase. It has been established that more than half of the β-amylase in barley is latent and in an inactive form which can be activated by the action of this type of protease; the effect is an almost trebling in the level of this carbohydrase in the mash.

Carbohydrase action reaches its maximum when the temperature of the mash is raised first to 65°C and then to 68°C, during which the starch becomes gelatinized and is converted by the combined action of the β-amylase from the barley, the added α-amylases and the contribution from the malt component of the mash. It has been rapidly established that, in addition to β-amylase, an amount of β-glucanase is required for the efficient control of the glucan gums that are extracted from the barley in the mash process. Many industrial fermentations for α-amylase production yield significant levels of a suitable β-glucanase for this purpose, and appropriate strains of *Bacillus subtilis* are cultivated for the preparation of a

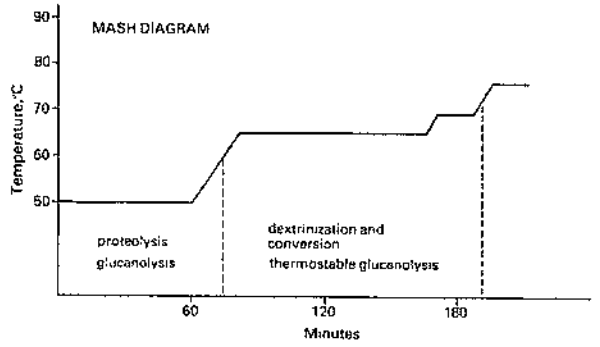


Figure 4.5.15 Mashing programme for barley brewing.

mixture of these two enzymes in appropriate proportions. Such enzyme preparations are added at the start of mashing, together with the proteinase, and it is considered that glucanase activity begins right away. Different producer strains of bacteria yield enzymes with small but useful differences of thermal stability and the more thermostable examples allow the maximum attack on glucan gums throughout the rising temperature of the programme. An example of this combined enzyme would be *Cereflo*® 200L from Novo, which would typically be used at a rate of 2.0 kilogrammes per tonne barley.

When α-amylase and β-glucanase are taken from separate sources, it is possible to adjust their proportions more precisely to the actual grist being used, and also to take advantage of both the alternative thermostable amylases, previously described in the section on adjunct cooking (*see* page 234), and the various β-glucanases. Selection of thermostable glucanases such as those from *Penicillium emersonii* in combination with thermostable amylases offers the opportunity of heating the mash more rapidly and so shortening the mashing cycle. There are additional benefits from the selection of more than one β-glucanase for the mash, as recent practice has demonstrated that there are small differences in the specificity of the various enzymes which combine to give a far higher rate of degradation of the most viscous materials. Although it is essential that the correct enzyme doses be established for the grist in question, some typical doses are indicated in Table 4.5.9. The dextrinization of the starch is then completed at the higher temperatures of the mash programme and the mash is finally raised to 78 °C before lautering.

TABLE 4.5.9
Suggested combinations of amylase and glucanase for barley brewing

	kg per tonne barley	*Combination*
Thermostable amylase (*Termamyl*® 60L Novo)	0.75–1.25	A
Conventional amylase (*Nervanase*® 180 ABM)	2.5–5.0	B
Thermostable glucanase (*Penicillium emersonii* ABM)	0.5–1.0	A/C
Standard glucanase (*Cereflo*® 200L Novo)	2.0–3.0	C
Standard fungal glucanase (*Finizym*® 200L Novo)	1.5–2.5	B/C

All-in-one enzyme systems. All-in-one enzyme systems are available in liquid form and allow a simple single addition of enzyme at the beginning of the mash. Since the internal ratios of the three declared enzyme activities – amylase, glucanase and proteinase – are fixed, one should not assume that they are necessarily optimally suited for all grists. In general, they contain a substantial excess of carbohydrases when used at the rates suggested and are therefore dosed on the basis of a correct protease level. This is not unreasonable, since there is little evidence of problems arising from the overuse of carbohydrases, while protein balance is far more critical for good beer production. It is also of importance from a practical monitoring aspect, since analysis of the nitrogen components of the mash (or the wort) is comparatively lengthy, whilst a simple iodine test will establish the satisfactory degradation of starch and an elementary filtration test for the removal of hindering gums takes only a few minutes. An example of the products available is *Ceremix*® from Novo, which is recommended to be dosed at 2.5–3.0 kilogrammes per tonne barley.

Extended saccharification. Only quite recently has extended saccharification in the mash tun vessels become a major target. This was previously carried out in the fermenter if an alteration in the amount of fermentable sugars became necessary. Growing concern regarding the labelling of finished beers has begun to alter the practice, however, despite the very severe economic penalties in the use rates of the selected enzymes. The combination of short contact times and adverse pH coupled with inactivation at the copper boiling mean that dramatically higher enzyme rates are required than would be the case at the fermenter (*see* page 249). However, where it is decided to increase fermentability at this stage it becomes a matter of choosing the enzymes according to the spectrum of fermentable sugars to be produced. Conventional amyloglucosidases and fungal amylases with maltogenic characteristics can be used by virtue of the same considerations already discussed for the saccharification of adjuncts (*see* pages 239 and 241). With a maximum of an hour in which to function, the dose rates must be much higher than those shown in Table 4.5.7 and can be as much as three times greater. As will be seen in the discussion on fermentation, these levels are anything up to 100 times greater than needed for the same conversion under more optimal conditions. Table 4.5.10 summarizes some typical barley brewing enzymes and their use rates.

Selection of enzyme rates to increase barley content. This can be a considerable task, and a foolproof rule has yet to emerge to aid the brewing chemist in making an accurate prediction. However, by

TABLE 4.5.10
Enzymes for barley brewing

Enzyme type	*Example product*	*Dose rate (kg per tonne barley)*
Bacterial neutral protease	*Neutrase*® 0.5L	1.2–1.5
Bacterial α-amylase	*BAN* 120L	2.0–3.0
Bacterial β-glucanase (with α activity)	*Cereflo*® 200L	2.0–3.0
Fungal β-glucanase	*Finizym*® 200L	1.5–2.5
Thermostable α-amylase	*Termamyl*® 60L	0.75–1.25
Thermostable fungal β-glucanase	*β-Glucanase* ABM	1.5–1.0
Combination products amylase, glucanase, protease	*Ceremix*®	2.5–3.0
Saccarification enhancing amyloglucosidase	*Amyloglucosidase* 150L	2.0–5.0 litres
Fungal amylase (maltogenic)	*Fungamyl*®	1.0–3.0

pooling the considerable practical experience gained in the past seven years, it is possible to draw some informative but tentative conclusions.

Laboratory extract measurements on the actual grist in the ratios at which it is to be used, yield data on the unaided performance of the intrinsic cereal enzymes present. If the results are plotted against typical doses of microbial enzymes that have been found suitable to achieve worts of high and reliable quality, a series of wedge-shaped plots of dose level emerge. Furthermore, the point at which the various enzymes become relevant can be observed. These data have been set out in Figure 4.5.16, which expresses the traditional extract in both Brewer's pounds per quarter and the Congress extract percentages. The approximate values for typical levels of barley at 25, 50 and 80 per cent of grist are indicated for reference.

It has been assumed that the most economical stage at which to add saccharification enzymes to improve fermentability will be at the fermenter, although by raising the base line dose as discussed earlier, these can be added to the mash. To establish the starting dose values for the left hand side of each 'wedge', it is practical to adopt the enzyme dose recommended (or experimentally determined) to give the maximal improvement of an all-malt mash using poor quality malt; this would be between 93 and 96 Brewer's pounds per quarter, or 75–77 per cent Congress extract.

The increasing use rates of each enzyme type then follow according to the figure, but on different gradients, so that amylases and glucanases increase eightfold, proteinase by one and a half times,

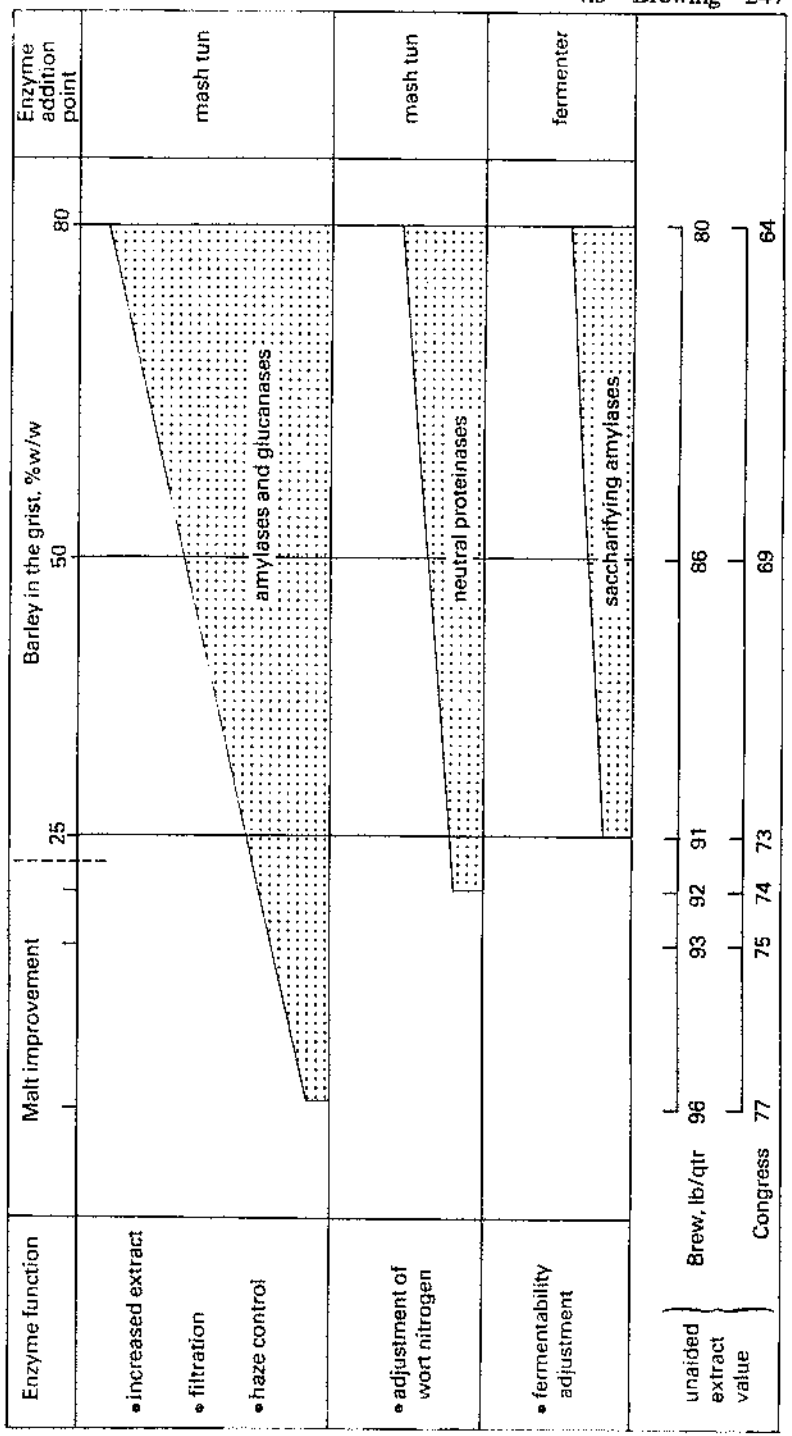


Figure 4.5.16 Selection of enzyme dose for increasing barley content.

and fermentability improvers by double as the barley rises to 80 per cent of the grist.

Economic considerations of barley brewing. Economic considerations alter each season according to the price and quality of both malting and feed barleys as well as the cost of good quality malts. It is not sensible, therefore, to declare hard and fast claims for the individual economic benefits that may be obtained, but the continued expansion of this practice in breweries around the world says much for the conviction that it contributes significantly to the economy of extract production. Indeed, the evaluation of the economic potential is most effective when related to the cost of extract, and the material from which it has been obtained. The choice between barley of 'feed' and 'malting' quality is a major factor in interpreting the benefits and a pooled survey of figures for Europe in 1980 is presented in Table 4.5.11.

10. Mash filtration

This stage is often described as the 'brewing bottle-neck', since it is here that the brewer has the first real indication of the degree of attack upon the cereal gums from his choice of grist and mash programme. Poor filtration is very time-consuming and is often reflected in the total extract and volume being lower than expected. The addition of the β-glucanase enzymes to the mash, both for malt improvement and barley brewing, can be very effective in restoring the desired filtration rate, as they are both more heat stable and of a wider specificity of attack than the enzymes found in the cereals. It is reported that for maximum benefit to filtration the mash should be supplemented with glucanases from more than one microbial species. Combinations of the enzymes from *Bacillus subtilis* and *Penicillium emersonii* have given excellent results when used together in ratios of from 70:30 though to 85:15 by activity (measured on a common assay basis). Fungal cellulases have also been shown to improve filtration, and this has been attributed to the wide range of glucanase-type activities that they exhibit. Furthermore, the cellulases give additional advantage in the processing of wheat

TABLE 4.5.11
Typical savings in extract cost when brewing with barley

% Extract from barley	*% Savings 'feed' barley*	*% Savings 'malting' barley*
30	18	14
50	27	22
70	36	30

in the mash because they often also contain hemicellulase (pentosanase) activity.

Examples of the effects of the presence of β-glucanase in the mash upon the filtration of worts are illustrated in Figure 4.5.17. From an economic standpoint a rate of enzyme is generally chosen that gives a 35–50 per cent improvement in filtration rate, with a viscosity reduction of between 15 and 20 per cent over the unaided mash, as determined in the brewery laboratory. As an example, this would indicate a rate of around 2.0 kilogrammes *Cereflo*® 200L (Novo) per tonne barley in the grist, and about 0.2 kilogrammes per tonne of grist for malt improvement. Table 4.5.12 compares the characteristics of wort produced by the application of the appropriate enzymes to a 50/50 barley malt mash in relation to that from a good malt and the unaided grist.

11. Enzymic treatments at the fermenter

Having produced a fermentation medium from a selection of raw materials processed by an appropriate method, some opportunities for adjustment still remain by treatment at the fermenter stage.

In the context of beer fermentation, several ways of expressing the proportion of the soluble wort solids actually converted to alcohol have been developed. For the purposes of this section the term 'real attenuation' will be used; this expresses the percentage soluble material converted to alcohol. The values vary depending

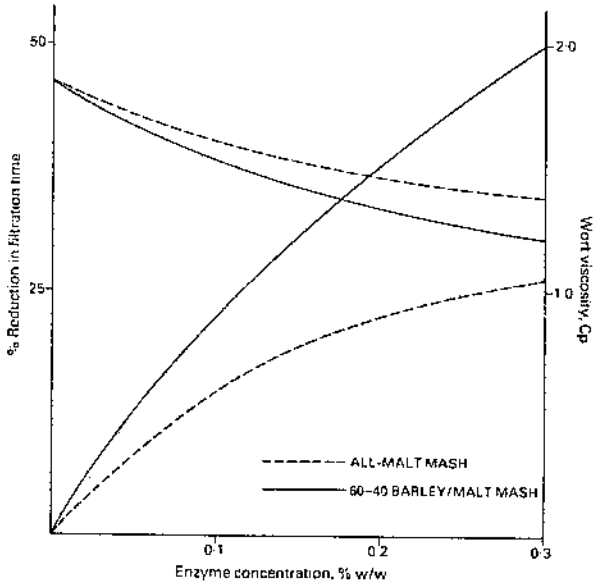


Figure 4.5.17 Effect of β-glucanase on mash filtration (*Cereflo*® 200L).

TABLE 4.5.12
Typical characteristics of barley brewing worts and beers

Raw materials	*Viscosity at 10°B Cp at 25°C*	*Extract °B*	*Reducing Sugar (g glucose per 100 ml)*	*Soluble N (mg per 100 ml)*
100% Quality malt	1.45	11.8	4.74	114
50/50 Barley/malt	1.71	11.3	4.12	67
50/50 Barley/malt + enzymes	1.38	11.7	4.63	108

on the composition of the grist, the manner of mashing and the yeast selected for the fermentation, together with the actual conditions of fermentation. For the production of most standard beers, real attenuation values lie in the range 58–68 per cent.

Adjustment of fermentation in standard beers. Only in exceptional circumstances is it necessary to make enzyme additions to the fermentations for these products, but there are situations in which it very significantly assists in recovery procedures. If heating at the mashing stage is inadequately controlled and too rapid or too high heating occurs, it is likely that a large part of the malt enzymes and any added enzymes will be destroyed before they have completed the conversion. Another situation is in the use of high syrup, or adjunct levels inadequately processed with saccharifying enzymes from microbial or malt sources that are unduly slow to ferment, and may even have a lower content of fermentable sugars than required for the target alcohol level in the beer.

The careful addition of small amounts of maltogenic fungal 'β'-amylase will reestablish the conversion of dextrins and facilitate correction of these defects. The very selective action of these enzymes on the unfermentable soluble dextrins is important in avoiding too high a degree of saccharification that would otherwise leave a beer with noticeably reduced 'body and mouth-feel'. Attack upon the α-1,6 branch points in the dextrins is minimal, and the main product is the splitting off of maltose units from peripheral chains, leaving the so-called 'limit dextrins' intact. These enzymes are described as self-limiting for this reason, and their use safeguards the brewer when he is making corrections in the production of his normal beers. Examples of such enzymes would be those from *Aspergillus oryzae* such as *Fungamyl*® 800L (Novo) or *Amylozyme*® 100L (ABM) used at doses of 2.0–20 grammes per hectolitre of wort being fermented. The maximum fermentability obtainable with these enzymes will give a real attenuation in the range 80–85 per cent as illustrated in Figure 4.5.18.

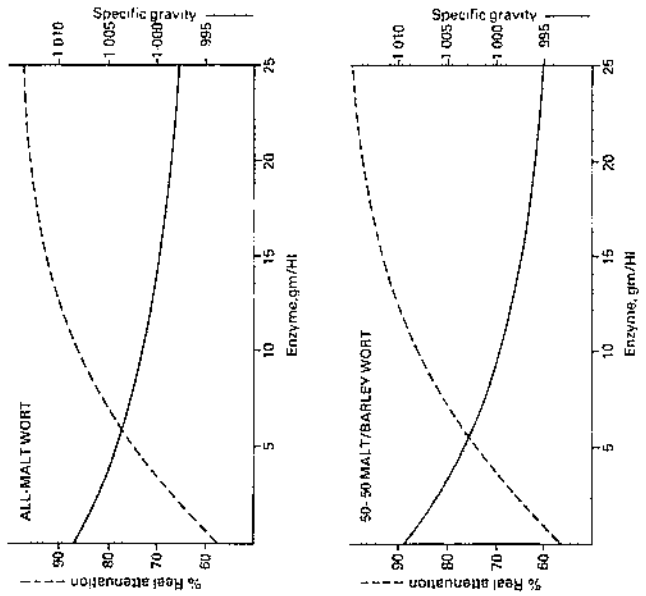


Figure 4.5.19 Increased fermentability by the addition of amyloglucosidase (*AMG* 150L; Novo).

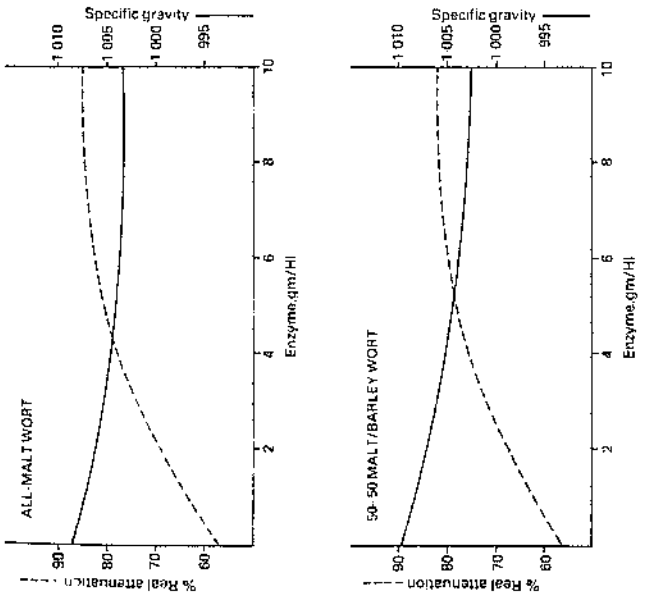


Figure 4.5.18 Increased fermentability by the addition of fungal 'β'-amylase (*Fungamyl*® 800L; Novo).

Maximum fermentability of worts. If the wort carbohydrates are degraded to the maximum amount of fermentable sugars (maltotriose, maltose and dextrose), such fermentation will proceed to the maximum only if a sufficiently long time is available. Although a rare requirement for normal brewing products, there is a demand in some countries for 'Brewer's alcohol' which is mixed into carbonated soft drinks to provide a beer content and a maximum alcohol content of 2 per cent. Addition of the amyloglucosidase enzymes from, for example, *Aspergillus niger* strains and variants, rapidly accelerates the saccharification by attack on both the α-1,6 branch points and the abundant α-1,4-linear chains of the dextrins in the wort. The equally rapid removal of the products of these reactions by the yeast fermentation provides an excellent thermodynamic drive for the enzyme to work steadily until all suitable substrates have been hydrolysed. Higher than normal alcohol levels, or standard alcohol levels can be achieved from either standard or lower wort gravities according to choice. Examples of these enzymes are *Amyloglucosidase* 150L (Novo) or *Ambazyme*® LE90 (ABM) added at the start of fermentation in doses from 5–25 grammes per hectolitre depending upon enzyme selection and the fermentation target. Some characteristics of this type of fermentation with all-malt and 50/50 malt/barley worts are given in Figure 4.5.19.

Production of 'diet' or 'lite' beers. The comparatively recent development of beer products with quite dramatically altered composition to meet the market interest for dietary control has taken the use of the two enzymes described in the previous two sections to a refined stage. Coupled with adjustments to the original gravity of the worts, selection of enzymes and enzyme ratios to suit the objectives is possible. These beers are described in many different ways, two distinct types emerge which make different claims.

The first type is 'low calorie' beer, which is subject to various regulatory definitions depending upon the country of sale. In the USA it is necessary that such a beer has only two-thirds of the calorific content of that considered to be representative of a normal beer. The second type is 'low carbohydrate' beer, which has a reduced level of unfermented carbohydrate residues. Such a reduction in carbohydrates does not necessarily confer 'low calorie' status, since the removed carbohydrate may well be present as alcohol, which has a higher calorific content than carbohydrate – 6.9 compared with 4.0 kilocalories per gramme for carbohydrate.

Low calorie beer. A normal American beer would be produced from around 11 per cent wort solubles with a fermentability of 64–68 per cent. The beer would have around 3.5 to 3.8 per cent alcohol, 2.9 to 3.1 per cent residual carbohydrates and 39 to 42 kilocalories per 100 grammes. For this wort to become a 'low calorie' beer, the energy content would have to fall to some 26 kilocalories per 100 grammes, which could not be achieved by simply raising the fermentability with amyloglucosidase enzymes, because at 85 per cent fermentability the beer would have 4.6 per cent alcohol, less than 1 per cent residual carbohydrates, but still around 38 kilocalories per 100 grammes. This problem has been solved by reducing the gravity of the wort to 7.5 per cent solids and fermenting to 85 per cent real attenuation. The target 26 kilocalories per 100 grammes can be reached at an alcohol value of 3.2 per cent. It will be appreciated that the use of high levels of unmalted adjunct in the mashing, or large amounts of syrups in the copper for these beers would seriously dilute the flavour and texture characteristics when fermenting from such low solids levels. These beers are usually produced from all-malt mashing, or from barley/malt mashes with barley substitution no greater than 50 per cent.

To obtain the desired fermentability, it has been found beneficial to use both amyloglucosidases and fungal 'β'-amylases together in proportions that will depend on several factors: (*i*) the intrinsic fermentability of the wort; (*ii*) the target fermentability; (*iii*) the desired retention of texture and 'body' in the beer; (*iv*) the desired alcoholic content; (*v*) consideration of the carry-through of enzymes in the beers. Some examples of the influence of different proportions of these two enzyme types are found in Figure 4.5.20, and the alcohol, carbohydrate and calorific values from the foregoing discussion are set out in Table 4.5.13.

TABLE 4.5.13

Examples of product characteristics for 'lite' beers

	% *Wort solids*	% *Real attenuation*	% *Beer alcohol (w/w)*	% *Residual carbohydrate*	*Kilocalories per 100 g*
Normal beer	11	65	3.6	3.0	40
Low calorie	7.5	65	2.4	1.95	26
	7.4	85	3.2	0.55	26
Low carbohydrate	7.0	85	3.0	0.5	25
	8.0	87	4.0	0.5	28
	8.0	74	3.0	1.5	29
	9.1	76	3.5	1.5	32

The thermal stability of the amyloglucosidases becomes relevant when carry-through to packaged beer is considered, for it is considered likely that at least 20 per cent of added enzyme will survive beer pasteurization. Conversely, the fungal 'β'-amylases are ther-

molabile and do not survive. With this in mind, there is a developing interest in the use of higher proportions of the latter enzymes, despite the fact that they need longer fermentation times. The resulting reduction in the level of amyloglucosidase is reflected in lower levels in the finished 'lite' beer product.

Low carbohydrate beer. As there are no clear cut normal values for the actual degree of reduction in carbohydrate values for this type of 'lite' beer, the demands on the production technique are lower. More significant constraints come in relation to the flavour, texture and alcoholic content, so that the adjustments are made by altering fermentability on different levels of wort solids. The mechanisms are much the same as for 'low calorie' beers. General experience shows that only when the residual carbohydrate is below one per cent does the 'low carbohydrate' beer also conform to 'low calorie' status. Slightly higher than normal alcohol levels are often found in this type of beer, which indicates that the wort gravities are maintained near to normal, probably to conserve some of the organoleptic qualities of a standard beer.

All the enzyme considerations mentioned in the production of 'low calorie' beers apply equally here, and in addition, there is some concern in many European breweries for the retention of good beer

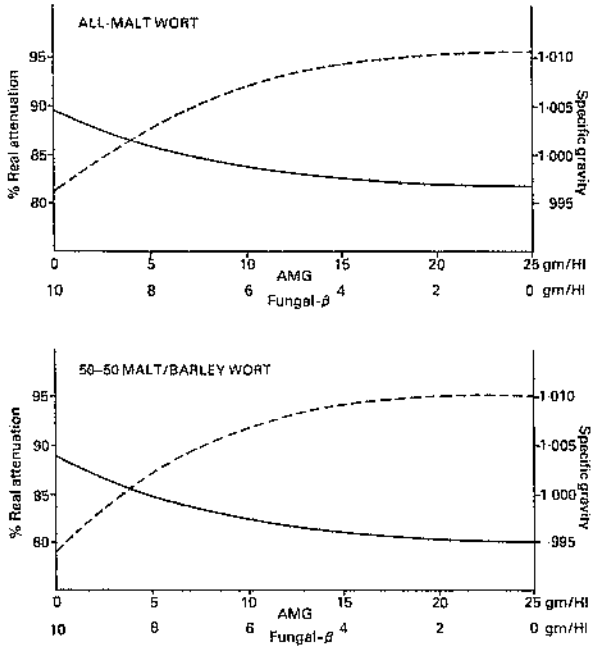


Figure 4.5.20 Increased fermentability by the addition of amyloglucosidase (*AMG* 150L; Novo) and fungal 'β'-amylase (*Fungamyl*® 800L; Novo).

foam (head) character. A number of amyloglucosidases particularly low in proteolytic side activities have been offered to meet the criticisms levelled against the standard enzymes. It remains to be seen if they will be entirely satisfactory, as detailed trials at some breweries are revealing several minor differences in the specificity of these proteases. These small differences can become large in regard to selective effects on head retention.

A third enzyme type for the conversion of wort carbohydrates has recently appeared; this has a potent debranching action on residual dextrins. Usually called 'pullulanase', but sometimes referred to as 'isoamylase', these enzymes may be of slightly differing action and become separately considered in the future. A combined tolerance towards fermentation pH levels, and a thermolability that prevents any carry-through risk, could make them very attractive. The cost may well delay rapid acceptance, although the heightened action in combination with the fungal 'β'-amylases keeps the use rates down to the range of 0.5–1.0 grammes per hectolitre of each.

Sweetening of finished beers (primings replacement). The addition of small amounts of sugar to sweeten certain dark beers is practised in many regions. The degradation of some of the residual unfermented dextrins in the beer after pasteurization can be used to generate the sweetness. Amyloglucosidase is generally used for this, but legislative considerations may mean that it will not be used so much in future. Addition before pasteurization requires that an allowance for the 70–80 per cent loss in the treatment is made in calculating the use rate of around 5 grammes per hectolitre, whereas only 1 to 2 grammes per hectolitre would be needed if the enzyme was added post-pasteurization.

Treatment of glucan hazes at the fermenter. The addition of fungal β-glucanases at this stage represents the last opportunity for the degradation of glucan gums that may give problems either at the beer filter or in the packaged product. As discussed on page 244, the choice of enzymes from *Aspergillus niger* or *Penicilium emersonii*, or a combination of both, provides a variety of substrate specificities that increases the potential for complete elimination of these haze-inducing polymers. The presence of glucans can be established on evidence from the filters or accelerated shelf tests on packaged beers and treated in following worts during fermentation. Typically, very low doses are required, 0.5 to 1.5 grammes per hectolitre being adequate. In some breweries it is the practice to delay the treatment until the wort reaches the maturation tanks, so that no further glucans can be contributed from dying yeast cells and the enzymes can provide a final solution to the problem.

12. Enzymatic chillproofing

This method of limiting the risk of haze in stored beers has been used for many years. It is becoming apparent that these hazes are complexes of several other substances including polyphenols and possibly carbohydrate fragments in addition to proteins, but also that the successful treatment of one component severely reduces the potential for the hazes to form. Papain, a protease extracted from the plant *Carica papaya*, is the preferred enzyme and is usually supplied to the brewing industry in a relatively pure liquid form. The mechanism of action of this enzyme in this situation is still not clear, but it is thought that it may initiate a coprecipitation of chill-haze proteins by a priming hydrolysis, in a manner similar to that of cheese rennets (*see* Chapter 4.6, 'Dairy').

Opinion varies as to the best stage for the addition of papain at the end of the beer production line, but it is agreed that it should be no earlier than the cold maturation tanks. This would provide the longest action time before pasteurization reduced the enzyme activity to very low levels. It is claimed for some preparations that the enzyme has been stabilized (it requires cysteine and a reducing environment for maximum stability) so that it will survive pasteurization. Such preparations are therefore recommended for addition immediately prior to pasteurization. Use rates vary widely with the activity and degree of stabilization of each preparation, but generally lie in the range 1.0 to 5.0 grammes of liquid products per hectolitre of beer treated.

Immobilized enzymic chillproofing. Recent interest in the application of enzymes immobilized into fixed packed columns of reactors has reached the brewing industry, and early trials with microbial enzymes prepared in this way have been encouraging. The mechanistic understanding for a satisfactory chillproofing action in only a short contact time as the beer flows through the reactors, and the absence of the enzyme protein to form a coprecipitate leave many questions to be answered. Perhaps the immobilized systems are operating on a totally different mechanism of haze formation. The single most attractive point of the use of immobilized enzymes is the total absence of the enzyme in the packaged product, but a further advantage is in the possible reduction in costs by having a system available for multiple re-use.

13. Related industries producing syrups, flavours and vinegar

Large areas of the process schemes and methods described in this chapter have been adopted, or independently arrived at, for the production of cereal extracts for other industrial applications. Syrups are prepared from both malt and barley grists to provide the flavour in the baking and confectionary industries as well as for the breakfast cereal market. Critical selection of the saccharification enzymes for the specific sugar spectrum and reaction with protein products forms the foundation of the use of enzymes to control flavouring (*see* Chapter 4.9, 'Flavouring and colouring'). Highly active enzyme syrups are also prepared as additional support to the malt complement, and emphasis is usually on the release of β-amylase by proteolytic addition as described on page 243. The preparation of barley syrups is largely for use by breweries with limited mashing capacity and uses the methods described on page 243, although generally higher enzyme use rates are applied to accelerate the processing. The use of thermostable enzymes for these syrups is not always recommended and must be related to knowledge of the market for the syrup. Where a diverse range of customer outlets is envisaged, it can be assumed that at least some of them will be for baked goods and other starch-containing products. If thermostable enzymes are to be used, it will be necessary to test the product from the syrup evaporators to confirm enzyme inactivation, so that starch degradation will not be a risk in the customer's product.

Non-alcohlic beer-like beverages are also prepared by enzymic conversion of a mash similar to that used for brewing standard beers, but with extra emphasis on the extraction of flavour and body-giving materials. The use of the maltogenic amylases is usually increased, together with additional proteolytic enzymes. High percentages of malt together with speciality malt-derived syrups form the basis of many of these products which are then either simply diluted and stabilized at low solids levels to be compatible with market acceptance or fermented as for the traditional product, the alcohol subsequently being stripped off under vacuum.

Where the regulations permit the wider use of cereals for the production of fermentation vinegars, most of the practices described for the cooking and saccharifying of adjuncts (*see* page 232) and barley brewing (*see* page 242) are utilized quite fully. Whilst some of the flavour character of the grain is required in all the brewed but not all distilled vinegars, the objective is largely to obtain a very high efficiency of alcohol production in the beer which is to form the feedstock for the acetification fermentation. Thus, for the vinegar brewer the use of saccharifying enzymes and their contribution both at mashing and during fermentation form an important part of his overall processing economy considerations. It should be noted that another factor is the use of the glucanases as regulators of haze development, a very important consideration for the vinegar producer, who markets a very clear product.

Where limitations on the grist put malt use at a premium, the use of enzymes for malt improvement is usually adopted for the best extraction at the mash tun, but enzymes are added less frequently to the saccharifying and fermentation systems.

14. Future developments in enzymes for brewing

Mashing and filtration enzymes. The development of a wider range of thermostable enzymes, in particular proteases and gum hydrolysing glucanases, can be expected. This will bring further reductions in the mashing times by reducing the need for a hold period at the initial lower temperature. This would also increase the possibility for the development of a downward temperature gradient programme, which is generally accepted as being a likely source of significant fuel economy at the brewery.

Further developments among the cellulases and pentosanases will increase the levels of wheat starch that can be processed successfully in single mashing systems. Fermentability adjustment at the mashing stage suggests a growing use for the saccharifying enzymes, and these should have higher thermostability than at present and must include the newer 'debranching' activities.

In the case of mashing capacity limitations, or the expansion of capacity without increased capital investment in mashing equipment, a rising grist to liquor ratio in the direction of the much-discussed 'high gravity brewing' can be expected. To simplify dosing and to provide economically acceptable extracts under these conditions, it is anticipated that higher potency enzyme preparations will be produced, and of higher priority still could be the development of more effective multi-enzyme products to enable one-shot additions to be made. Increasing automation in modern brewhouses should be part of the planning of enzyme treatments, too, and single dose systems will fit most readily into automated programming.

Adjunct processing enzymes. An increased use of saccharifying enzymes to prepare highly fermentable worts to add at the copper, rather than at the mashing stage, can be expected to reduce extract costs. By this means, the equivalent of starch syrups can be provided at the brewery in areas where they are not readily available from starch processing industry.

New processing systems for reduced energy at adjunct cooking have high priority. It has been calculated that, by using the continuous cooking systems of the starch industry together with a downward temperature programme for conversion and saccharification, a further reduction in fuel of 20–25 per cent can be established (*see* Chapter 4.15, 'Starch').

Fermentation enzymes. In the efforts to limit the carry-through of enzymes to finished products, it is anticipated that saccharifying and debranching enzymes will be prepared in thermolabile forms, but otherwise retain the same activities as used for precopper saccharification.

Probably the most glamorous developments will come from brewing research in the genetic engineering of yeasts. One target is the production of brewing yeasts with their own complement of saccharifying enzyme activity, and early results show that the concept is entirely feasible. The distilling industry, in particular the fuel alcohol side, has an equally large interest for similar yeasts, although opinion differs about the temperature of fermentation considered optimal, and through this route we can see more chance of covering research costs. Between them, the two industries will almost certainly fund the development of such yeasts and then incorporate them into regular production applications once they are regulatorily approved.

Chillproofing enzymes. As more detailed knowledge of the nature of the chill-haze compounds and the mechanism of the haze dissolution from the beer is gained, it is likely that more specific enzymes can be selected to attack one or more of the components to prevent haze formation. The survival of papain into the beer is considered part of its chillproofing function and future work will be directed to establishing if this is the case. If not, then the immobilization of papain together with its stabilizing chemistry will be a commercial target. Similarly, the immobilization of other proteases, and possibly tanninase and polyphenolases from the research into cellulose utilization (which has a high priority for removal of lignin), will be used to develop alternative chillproofing systems.

Further reading

Proc. Congr. European Brewing Convention: 13th Congress, Estoril (1971); 14th Congress, Salzburg (1973); 15th Congress, Nice (1975); 16th Congress, Amsterdam (1977); 17th Congress, Berlin (1979); 18th Congress, Copenhagen (1981) (Elsevier, Amsterdam).

Proc. a. Meet. Am. Soc. Brewing Chemists.

Hough, J. S., Briggs, D. E. & Stevens, R. *Malting and Brewing Science* (Chapman & Hall, London, 1971).

Harborne, J. B. & van Sumere, C. F. *The Chemistry and Biochemistry of Plant Proteins* (Academic, New York, 1975).

Rose, A. H. *Economic Microbiology* Vol. 1 (Academic, New York, 1977).

INDUSTRIAL ENZYMOLOGY

The Application of Enzymes in Industry

TONY GODFREY
and
JON REICHELT

M
The Nature Press

Published in the United Kingdom by
MACMILLAN PUBLISHERS LTD (Journals Division), 1983
Distributed by Globe Book Services Ltd
Canada Road, Byfleet, Surrey, KT14 7JL, England

ISBN 0 333 32354 8

Published in the USA and Canada by
THE NATURE PRESS, 1983
15 East 26th Street, New York, NY 10010

Library of Congress Cataloging in Publication Data
Main entry under title:
Industrial enzymology.
Includes index.
1. Enzymes – Industrial applications. I. Godfrey, Tony, 1939– . II. Reichelt, Jon, 1949–
TP248.E51516 1983 660'.63 82–14461
ISBN 0–943818–00–1

Printed in Great Britain

FRUIT JUICE

W. Janda

1. Introduction

The first enzymes to be used by the fruit juice industry were the pectolytic enzymes for the clarification of apple juice. They were introduced in 1930 simultaneously by Z. J. Kertesz (1930) in the USA and by A. Mehlitz in Germany. Since then, fruit juice processing has developed into a major, highly technological industry, covering not only core fruits but also stone fruits, citrus and tropical fruits, berries, grapes and even vegetables. The functions of pectolytic enzymes have become more specialized and other enzymes such as amylases and cellulases also now form an integral part of today's fruit juice technology.

2. The different enzymes

Pectolytic enzymes. The most common enzymes in use in the fruit juice industry are still the pectinases. Their substrate, pectin, is an essential structural component of fruits where, with hemicellulose, it binds single cells to form a tissue. In the immature fruit, pectin is mainly insoluble and becomes partially soluble as the fruit matures and becomes softer. During the juice processing, when the plant tissue is disintegrated, some of the pectins go into solution, some become saturated with juice and some remain on the cell walls. Pectins from some fruits can hamper the processing or lower the quality of the juice, so that in most cases it is desirable at least to modify the pectins or even to break them down completely.

Amylases. In the late sixties and early seventies, in addition to fruits grown especially for their juice, the apple and pear juice industry began to process table fruits from cold storage warehouses. This fruit is normally picked before it is completely mature to ensure its firmness, and is then ripened under a controlled atmosphere in the cold store. Because it is not completely mature, it still contains starch, and this becomes gelatinized during juice processing and may lead to problems with filtration or haze formation. The use of amylases, especially amyloglucosidases, has become a routine way of overcoming such problems.

Cellulases. Cellulases, either as a single preparation or as part of the pectolytic enzyme preparation, are useful tools for speeding up the extraction of colour from the fruit or for the total liquefaction of

the plant tissue. For example, the colour in blackcurrants is a valuable component of the fruit, but it is located in the skin cells, which are harder to disintegrate than the juice-containing cells in the fruit flesh. Extraction is also hampered by the poor permeability of the cell walls and membrane. This can be accelerated by improving the cell wall and membrane permeability either by performing plasmolysis at temperatures of 60 °C and higher or by the use of cellulases in addition to pectolytic enzymes at temperatures of 50 °C or lower.

When used with the total liquefaction method, the effect of the cellulases is much more pronounced, so that the plant tissue becomes completely macerated and the separation of the liquid and solid parts of the fruit can be carried out by filtration, centrifugation or static decantation instead of by pressing.

3. The use of enzymes in the production of different fruit juices

Core fruits, stone fruits and berries. Figure 4.10.1 is a typical flow sheet of a fruit juice processing line. After the fruits have been washed and sorted, and where applicable destoned and destemmed, they are disintegrated in a mill and heated to the temperature required for enzyme-prepress treatment. This is typically performed with *Pectinex® 3XL* (Swiss Ferment Co.) at 3–20 grammes per 100 kilogrammes fruit, and *Celluclast® 2.0L* (Swiss Ferment Co.) at 0.2–2.0 grammes per 100 kilogrammes fruit). The optimal temperature for enzyme-prepress treatment of core fruits is 30 °C, while for stone fruits and berries the optimal temperature is 50 °C when cellulases are used to improve colour extraction, and 60–65 °C when extraction is performed by plasmolysis.

The prepress treatment also helps to break down the insoluble pectin, which occurs as small, slimy jelly-like particles. These

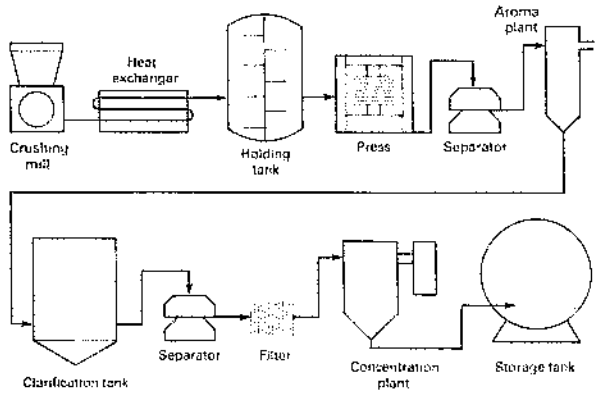


Figure 4.10.1 Fruit juice processing line.

hamper the extraction of the juice in two ways. First, they become saturated with juice, which is then not available for pressing, and second, the particles block the small drainage canals in the pomace through which the juice must run.

Prepressing can be considered to be complete when the juice reaches the desired colour level and when its viscosity has returned to at least its initial value, or even lower (at the beginning of the enzyme reaction, the viscosity increases due to the solubilization of the insoluble pectin). The core fruit crush should not be heated beyond 30 °C because this can destroy its physical structure, which is essential for the pressing operation.

After the mash has been pressed, the aroma of the juices is stripped off and the juice goes into the clarification tank. Juices from core fruits are passed through a centrifuge before they go to the aroma plant in order to separate the main part of the unjellified starch. The remaining part of the starch (some five per cent) is jellified in the aroma plant. Depectinization, breakdown of starch and clarification of the juices can be performed at 20–25 °C or 45–50 °C, typically by the addition of *Pectinex® 3XL* at 1.5–3.0 grammes and *Amylase AG 150L* at 0.5–2.0 grammes per hectolitre (Swiss Ferment Co.). The higher temperature gives an advantage due to the fact that enzymes in general, including the pectolytic and amylolytic enzymes, are more active at higher temperatures. The temperature range 25–45 °C must be avoided, however, because it creates ideal conditions for the growth of microorganisms, especially yeasts.

Depectinization has two effects: it causes coagulation of the cloud, which is stabilized by insoluble pectin, and it breaks down the viscosity-causing soluble pectin. As mentioned above, when juices from core fruits are being processed, the remaining five per cent of the jellified starch must be broken down at this stage by means of an amyloglucosidase.

After the fining of the juice, which must not be performed before the juice is completely depectinized and free of starch and dextrins, the clear juice is separated by means of static decantation, centrifugation and/or filtration from the fining precipitate, and either concentrated or pasteurized and bottled at single strength.

In the USA, where most of the fruit juice is bottled at single strength and where the consumer will accept a slightly hazy juice, the juice from the press is brought directly to the clarification tank, and the pectins are broken down only as far as required for the juice to be filtered. Fining is normally performed with gelatine alone, at the same time as the enzyme reaction. The fining precipitate is then filtered and the juice is pasteurized and bottled.

Several years ago, a new process, originating from the sugar beet industry, was introduced, which replaces the operation unit pressing by countercurrent extraction (Schobinger, 1978). Juices obtained from this process (presently mainly apple juice) normally have lower pectin levels than those obtained by pressing, but are depectinized and clarified in the same way as are pressed juices.

A future trend in this branch of the fruit juice industry will be the total liquefaction of fruits. As mentioned above, the objective of this method is to macerate/disintegrate the fruit tissue with pectolytic, cellulolytic and other enzymes to such an extent, that the liquid and solid parts of the fruit can be separated by methods other than pressing. The final product would then contain almost everything originally present in the whole fruit.

Grape juice. White grape juice and its concentrate are produced by similar methods to those used for core fruits. The grapes are crushed and destemmed, and the crush is prepress-treated with pectolytic enzymes in the dejuicer to increase the yield (typically with *Pectinex® 3XL* at 1.6–5.0 grammes per 100 kilogrammes fruit). After the free-running juice has been drained off, the remaining crush is pressed. The juice obtained is then depectinized in clarification tanks, sometimes fined, stabilized, centrifuged and/or filtered, pasteurized and concentrated or bottled single strength.

Red grapes, especially concord grapes, have a high pectin content, which makes the crush slippery and difficult to press. Furthermore, an objective of the processing is to extract almost all the colour from the skin. Therefore, the method is somewhat different from that used to obtain white grape juice. The crushed and destemmed red grapes are brought to 60–65 °C, where plasmolysis is performed to accelerate colour extraction and where the crush is treated with pectolytic enzymes for approximately half an hour to eliminate the slipperiness and increase the yield. In some cases, the crush is heated beforehand to 80–82 °C to destroy the fruit oxidases and so prevent oxidation and loss of colour.

The free-running juice is then drained off and the remaining crush pressed. Free run and press juice is depectinized with pectolytic enzymes, gently fined if necessary, stabilized, centrifuged and/or filtered and either concentrated or bottled at single strength.

Citrus industry. The four main applications of enzymes in the citrus industry today are: (*i*) in pulp wash; (*ii*) to lower the viscosity of orange juice concentrate; (*iii*) in the preparation of natural cloudifiers; (*iv*) for the clarification of lemon juice.

(*i*) Pulp wash. During the processing of citrus fruits, juice, peel and pulp are obtained. The pulp, which forms roughly a quarter of the total fruit, contains considerable amounts of juice which cannot

be easily extracted by pressing. Methods have therefore been developed to obtain this juice using a three to five step countercurrent extraction (pulp wash) of the pulp with water. Pectolytic enzymes are used in this process to treat the pulp before the extraction and so increase the washable solids, and also to lower the viscosity of the pulp wash juice, so that a concentrate of 60 °BX can be obtained without risk of jellification.

The treatment of the pulp before extraction is done either in a continuous or a batch process. Pectolytic enzymes are thoroughly mixed with the pulp and allowed to act for about 30 minutes to break down the insoluble part of the pectin and release the trapped juice (typically *Pectinex® 3XL* (Swiss Ferment Co.) at 2.5–5.0 grammes per 100 kilogrammes pulp is used). In the treatment of the pulp washing liquid, a very limited breakdown of the pectins is carried out, just sufficient to reduce the amount of soluble pectins and so lower the viscosity without attacking the insoluble pectin fraction which maintains the stability of the cloud.

(*ii*) Lowering the viscosity of orange juice concentrate. Orange juice, prepared from certain types of fruit, can have a high viscosity and so may undergo jellification if concentrated up to 65 °BX. As with the pulp wash liquid, these problems are overcome by the use of small amounts of pectolytic enzymes (typically *Pectinex® 3XL* (Swiss Ferment Co.) at 1.6–3.5 grammes per hectolitre juice). It should be mentioned, however, that this application of enzymes is illegal in some countries, as, for example, in the USA.

(*iii*) Preparation of natural cloudifier. Because the use of brominated oils and artificial cloudifiers in citrus beverages is prohibited in several countries, the demand for suitable, natural cloudifiers, originating from the citrus fruit itself, has increased considerably. Among the different processes available today for the production of peel extract concentrates, the most effective are those using pectolytic enzymes. A typical flow sheet for such a process is given in Figure 4.10.2.

The raw material, citrus peel, to which pulp and rags may also be added, is ground to an average particle size of 3–5 mm. It is then mixed with water at 1:1 to 1:1.5 on a weight basis, heated to 95 °C to destroy the fruit pectin esterase and cooled to 50 °C. (If the raw material has a low content of pectin esterase, it need only be heated to 50 °C.) Pectolytic enzymes are added and allowed to act either batchwise or continuously for a half hour to an hour (typically *Pectinex® 3XL* at 3.5 grammes per 100 kilogrammes peel is used). During that time, the enzymes bring about a kind of maceration of the peel and release cloudy material such as cellulose, pectins, hemicellulose and cell organelles into the liquid. This

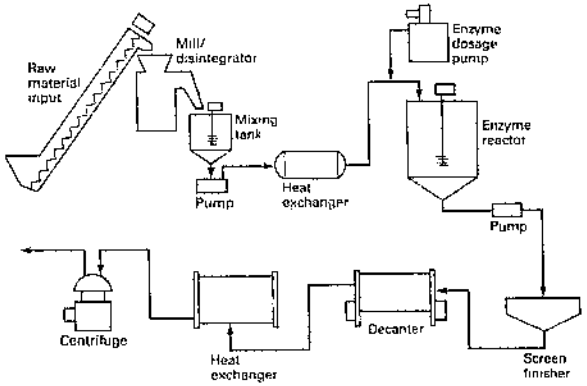


Figure 4.10.2 Plant for citrus peel extraction.

liquid is then separated from the solids, pasteurized and concentrated.

(*iv*) Lemon juice clarification. The traditional way of clarifying lemon juice involves protecting the juice from microbial spoilage by using some 1000–2000 milligrammes of sulphur dioxide per litre to allow self-clarification to occur. Depending on the pH of the juice and the temperature, this can take some 4 to 16 weeks. Such a process therefore requires high storage capacities.

Recently, pectolytic enzymes showing improved activity at very low pH values (2.2–2.8) have been available which modify the pectin in lemon juice within 3 hours at ambient temperature to such an extent that the cloud can be coagulated and precipitated with an agent such as silica sol (typically *Pectinex® 3XL* at 11–20 grammes per hectolitre is added). The whole process, from the fresh fruit to the final concentrate, is shortened to some 6 hours.

Future developments in the citrus industry seem to lie with debittering enzymes. Citrus bitter principles like naringin and limonin occur in certain geographical areas, and in certain fruits make it difficult to produce a juice with a pleasant but not too bitter taste.

Naringin, one of the bitter principles of grapefruits, can be converted almost completely enzymatically to its aglycone. This is not possible with limonin, which forms the main bitter substance in grapefruits and oranges. Several groups of scientists, however, are now working on the modification of limonin and its precursor, and may soon come up with an industrially feasible solution to the problem.

References

Kertesz, Z. J. *Bull. N.Y. St. agric. Exp. Stn.* **589** (1930).
Mehlitz, A. *Biochem. Z.* **221,** 217–31 (1930).
Schobinger *Frucht- und Gemüsesafte* (Ulmer, Stuttgart, 1978).

INDUSTRIAL ENZYMOLOGY

The Application of Enzymes in Industry

TONY GODFREY
and
JON REICHELT

M
The Nature Press

Published in the United Kingdom by
MACMILLAN PUBLISHERS LTD (Journals Division), 1983
Distributed by Globe Book Services Ltd
Canada Road, Byfleet, Surrey, KT14 7JL, England

ISBN 0 333 32354 8

Published in the USA and Canada by
THE NATURE PRESS, 1983
15 East 26th Street, New York, NY 10010

Library of Congress Cataloging in Publication Data
Main entry under title:
Industrial enzymology.
Includes index.
1. Enzymes – Industrial applications. I. Godfrey, Tony, 1939– . II. Reichelt, Jon, 1949–
TP248.E51516 1983 660'.63 82–14461
ISBN 0–943818–00–1

Printed in Great Britain

INDUSTRIAL APPLICATIONS

Chapter 4.1

ALCOHOL – POTABLE

P. B. Poulson

1. Introduction

Potable alcohol is the term used for all distilled spirits (ethanol content higher than around 20 per cent) intended for human consumption. The term covers an enormous number of different spirits produced by many different processes, and consequently this chapter aims only at covering the most important parts of potable alcohol production.

Potable alcohol has been produced industrially as well as domestically for many hundreds of years. In fact, it is one of the oldest industries. The industry presumably grew from chance fermentation of sugar-containing juices followed by natural distillation by the sun. Examples like the fermentation of cactus juice (which has now been developed as Tequila a.o. (Mexico)) and the fermentation of palm juice (e.g. Ogogoro (Nigeria)) can be mentioned. The industry later became more sophisticated with the finding that barley could be transformed to malt and thereby be used in the processing of starch-containing crops to alcohol (e.g. whisky (Scotland)). A further improvement was made in China and Japan where special microorganisms were grown on cooked rice in order to produce starch-fermentable agents (koji). Around 1890, the enzymes present in koji were extracted and concentrated by Takamine and sold as takadiastase. This was the beginning of the modern enzyme industry. Thus it can be seen that the history of the development of the potable alcohol industry is in fact also the history of the growing importance of industrial enzymes.

When sugar-containing crops are to be used as raw materials for the production of spirits, there is only a need for ethanol-producing agents. However, when starch materials are to be used, it is essential that the starch is hydrolysed to fermentable sugars. In some areas, this was done by means of enzymes of vegetable origin (malt), and in others, by means of enzymes of microbial origin (koji). Both methods have survived, although there has been a tendency – especially during the past 20 years – towards the replacement of

malt (and originally koji) by industrially produced enzymes (from microbes). The reason for this replacement has been partly economical and partly the need for a product of consistent quality.

2. Raw materials

Many varieties and mixtures of sugar-containing and starch-containing raw materials are used in the production of potable alcohol. Tables 4.1.1 and 4.1.2 give some examples, while Table 4.1.3 shows the relative proportions of the raw materials used.

TABLE 4.1.1
Examples of potable alcohol produced from sugar-containing raw materials

Raw materials	*Products*
Molasses (sugar cane)	Caribbean rum Brazilian cachaça
Wine (grapes)	Cognac Pisco (Peru)
Agave azul tequilana	Tequila (Mexico)
Cherry	Kirsch (Switzerland)
Pear	Pear brandy (Switzerland)
Plums	Slivovice (Balkans)
Palm juice	Ogogoro (Nigeria)

TABLE 4.1.2
Examples of potable alcohol produced from starch-containing raw materials

Raw materials	*Products*
Barley	Whisky
Maize and rye	Bourbon whiskey
Potatoes and barley	Aquavit
Potatoes, rye, wheat (etc)	Vodka
Rice	Chinese brandies

3. Applied cultures and enzymes

Starch-hydrolysing agents include the following.

Malt. Malt is germinated barley. During the germination, enzymes are formed or activated. Enzymes of special interest for brewing processes are:

Starch-hydrolysing enzymes (α- and β-amylases)
Protein-hydrolysing enzymes (proteases, peptidases)
Hemicellulose-hydrolysing enzymes (cytases)
Phytin-hydrolysing enzymes (phytases)

When barley is to be transformed to malt, it must be made to germinate. This requires a water content of 42–46 per cent in the barley. The malt process is initiated by soaking (steeping) the barley in water for two or three days, and then allowing it to germinate, a process which takes six or seven days. The temperature (10–22 °C) during this step is very important in determining which type of malt is produced. At this stage, the malt is called green malt, and the final step is to dry to dry the green malt. The drying step causes:

Reduction in the water content from around 45 per cent to 1.5–4 per cent.
Cessation of germination and digestion.
The formation of colouring and aromatic compounds.

On average, 100 kilogrammes barley will give around 80 kilogrammes malt.

Koji. The manufacture of koji can be performed as follows. Dehusked brown rice (unpolished) is pounded briefly with a wooden pestle in order to scratch the surface of the outer epidermis. The rice is then washed thoroughly, soaked in water overnight and cooked in live steam for about an hour. It is then taken out of the steamer, cooled, and mixed manually and evenly with burnt wood ash (two per cent by weight). When the temperature of the mixture is low enough so as not to be harmful to the microbes, the powdery koji seed, which is carefully preserved by successive transplantations, is sprinkled over the rice and rubbed with the fingers so as to distribute the spores and to bring them into contact with the surface of the rice. The whole mass is incubated overnight in a warm koji chamber until the temperature of the mass has risen to 35 °C. The mass is then divided among small shallow wooden trays which are piled up in the chamber. The temperature and humidity are controlled by changing the type of piling and by opening the ceiling window.

The propagation of mould as well as the abundant formation of spores can be completed within five or six days, after which the trays are taken out into the open air and each tray is covered with thin paper and exposed to direct sunlight for one day. The mass is then carefully dried again at 40 °C in an indirectly heated drying chamber, and the final product is wrapped in a paper bag and stored.

Microbial enzymes. The market share of the abovementioned enzyme products has been declining, especially over the past ten years, as an increasing number of distilleries have decided to use industrially produced microbial enzyme products. The reasons for this have been both economical and technical, the technical reasons being that the microbial enzyme products have a known standardized activity and are much more concentrated, and thus require less handling.

The microbial enzymes are produced by special, carefully optimized mutants in very large vessels – typically 100–200 cubic metres. The enzyme products on the market are either single enzyme products (α-amylase, β-amylase, glucoamylase, protease etc) or enzyme mixture products.

Ethanol-producing organisms. Whereas starch hydrolysis has almost become an art, the selection of ethanol-producing organisms has been rather casual. For example, in the fermentation of grapes, many wineries are quite satisfied with yeasts of natural origin. Generally speaking, yeasts, such as *Saccharomyces cerevisiae*, *Saccharomyces carlsbergensis* and *Schizzosaccharomyces* types, are used in temperate climates for the production of alcohol, while in tropical countries, bacteria (e.g. *Xymomonas mobilis*), are used.

4. Processes involved

Main processes in distilling

The manufacture of alcohol from starch-containing raw materials is based on the following main processes (*see* Figure 4.1.1).

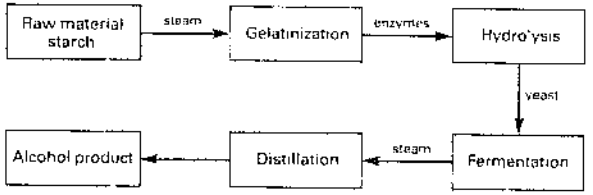


Figure 4.1.1 Main processes

Gelatinization. The dissolution of the raw materials into a mash by steam cooking to make the starch available for enzymatic attack. Normally, a concentration of 15–20 per cent starch in the mash is aimed at.

Hydrolysis. The breakdown of the dissolved starch to fermentable sugars by means of enzymes.

Fermentation. Conversion of the sugars to alcohol by the action of yeast.

Distillation. Separation and purification of the alcohol.

The enzymatic hydrolysis consists of two stages:

Liquefaction. The gelatinized starch is broken down into short molecule fragments (dextrins) by means of α-amylase, resulting in a

rapid reduction in mash viscosity.

Saccharification. The dextrins formed during liquefaction are further hydrolysed to fermentable sugar (glucose) by means of glucoamylase.

American batch process. In the American batch process (*see* Figure 4.1.2) milled corn is slurried with water and fed into the cooker at a concentration of about 25 per cent, where it is kept under agitation. A thermostable, liquefying bacterial α-amylase (e.g. Novo *Termamyl*® 60L at 0.15–0.3 kilogrammes per tonne) is added, and the temperature of the 'mash' is gradually increased by the injection of live steam to about 150°C. During cooking, the starch becomes gelatinized and the mash viscosity increases. The liquefying α-amylase partially hydrolyses the gelatinized starch and reduces the mash viscosity sufficiently to allow agitation to continue. This is referred to as the pre-liquefaction stage.

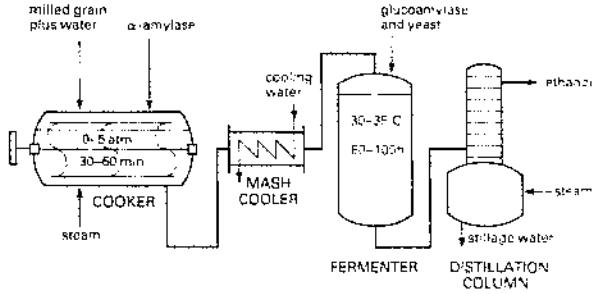


Figure 4.1.2 American batch process

As heating continues up to 150°C, the remainder of the starch becomes gelatinized, but the enzyme has by now been inactivated. However, the starch granules have been completely disrupted and the starch is fully susceptible to enzyme attack. The mash is cooled and further bacterial α-amylase added (e.g. Novo *Termamyl*® 60L at 0.35–0.7 kilogrammes per tonne). The enzyme partially hydrolyses the gelatinized starch and reduces the viscosity of the cooling mash and prevents starch retrogradation. This is referred to as post-liquefaction. After the post-liquefaction stage, the mash is cooled to about 60°C and the saccharifying enzyme (glucoamylase) added (e.g. Novo *SAN* 150L at 1.5–2.0 litres per tonne). Yeast is added when the temperature has been lowered to about 30°C. The glucoamylase converts the partially hydrolysed starch to dextrose which is fermented to ethanol. Simultaneous saccharification and fermentation continues for 60–100 hours, after which the ethanol is distilled off.

German batch process. In the German batch process (*see* Figure 4.1.3), the raw material is gelatinized without previous milling by

cooking with live steam in a Henze cooker. No addition of enzyme or mechanical agitation are necessary in the cooking stage. The cooked mash is blown through a strainer valve into the mash tub where the liquefaction takes place according to one of two procedures:

(*i*) High temperature liquefaction: The blow-down is carried out within the shortest possible time, after which the mash is cooled in the mash tub to 80°C. At 80°C, α-amylase (typically 0.15–0.6 kilogrammes Novo *Termamy*® 60L per tonne of starch) is added and the temperature is maintained for 20 minutes before further cooling.

(*ii*) Low temperature liquefaction: Before blow-down of the cooker, the mash tub is filled with cold water, sufficient to cover the lowest part of the cooling coil, and α-amylase is added (e.g. Novo *Fungamyl*®800L at 0.1–0.2 kilogrammes per tonne). The mash is blown down into the mash tub under agitation and cooling at such a rate that a temperature of between 55 and 60°C is maintained in the mash tub. When the blow-down is complete, the mash may be further cooled to the fermentation temperature.

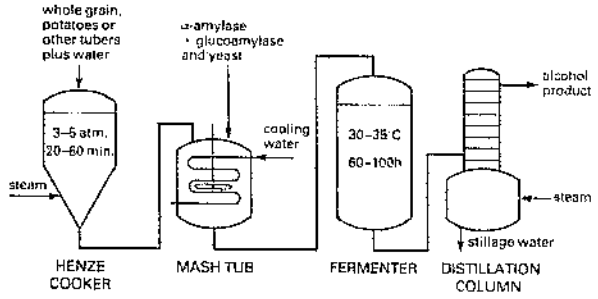


Figure 4.1.3 German batch process

Saccharification. Glucoamylase (e.g. 1.1–2 litres Novo *SAN* 150L per tonne of starch) is added to the liquefied mash at 60°C or lower. This will bring about complete saccharification within the normal period of fermentation. Yeast is added to the mash after cooling to 30°C or lower.

Continuous cooking. Cooking and liquefaction may be carried out continuously, thus giving better process control and more efficient use of equipment. Such a process is shown in Figure 4.1.4. A milled corn slurry to which a thermostable bacterial α-amylase has been added is heated by direct steam injection to about 150°C in a jet cooker. The mash is then flash-cooled to 80–90°C and the second addition of α–amylase made. The cooked mash is held at this temperature for 30–60 minutes to complete liquefaction before it is

transferred to the saccharification/fermentation tanks. (Continuous cooking followed by continuous fermentation has also been described by Rosén (1978).) Enzyme dosages are the same as for the American batch process.

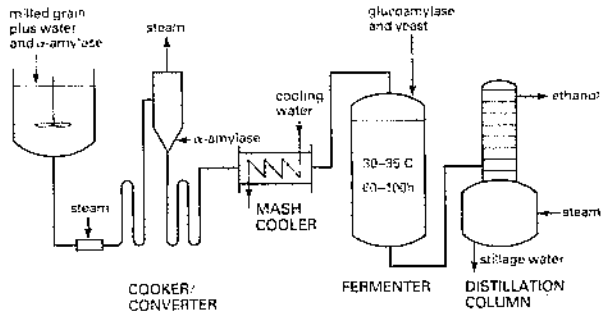


Figure 4.1.4 Continuous cooking process

Newer process layouts and future process developments. What will happen to the technical processes used in the potable alcohol industry in the coming years? I think that the processes will be modified, especially with respect to energy consumption and more controlled processing of the raw materials. One should bear in mind that the potable alcohol industry is a very old one where traditions are very important. It is not possible to change raw materials and processing without years of testing in order to ensure that the quality of the products has not changed.

In the traditional batch processing method, energy consumption was about 17–24 megajoules per litre ethanol (Hagen, 1981). Of this amount, 7–8 megajoules per litre were consumed in the cooking of the raw materials, the rest during the distillation. With modern continuous processing methods, the indications seem to be that the energy consumed can be decreased to 6–9 megajoules per litre (1–2 megajoules per litre during cooking and 5–7 megajoules per litre during distillation) without changing the quality of the product. This saving is obtained by increasing the temperature at which the cooker feed is preheated from 20°C to 60°C using recycled surplus process energy, increasing the mash concentration from 20 to 30 per cent dry substance and reducing the cooking temperature from 150°C to about 100°C (the yield loss is limited to a very few per cent, dependent on grain quality and particle size).

Recently, work in Japan (Suntory) has shown that it is possible to reduce the processing temperature to 35°C when using microbial enzymes and a holding time of up to five days. Presumably, further optimization will be reported within this field in the coming years.

The continuous development which is taking place is also important in that it allows industrial enzymes of microbial origin to replace at least part of the expensive and troublesome use of malt and koji. Furthermore, more and more enzymes are being marketed, enzymes which can perform increasingly more specialized reactions and make the substitution for malt and koji easier so that commercial benefits can be obtained without changing the quality of the potable alcohol.

5. Production size (Schrøder)

When trying to determine the basic production of potable alcohol (*see* Table 4.1.4), one runs into a number of statistical problems in addition to the normal statistical uncertainties involved.

One problem is that, owing to the fact that potable alcohol is taxable, it is tempting for manufacturers to give too low production figures, which may be a significant factor of uncertainty in some countries. A more serious and frequent problem is the fact that industrial alcohol is sometimes included in distilled alcoholic beverages. This category is the most heterogeneous in other ways too, for it may include beverages with an alcohol content varying from a few per cent to 80 per cent. This heterogeneity may become more marked as new types of mixed drinks are introduced based on distilled alcohol but with a very low alcohol content.

Most countries, however, report production in consumption strength which is assumed to be 40 per cent by volume, on an average. This assumption is in accordance with the practice of the Produktschap Voor Gedestilleerde Dranken (Netherlands).

TABLE 4.1.3

Relative proportions of raw materials used for production of potable alcohol

	100% ethanol ($\times 10^6$ hectolitres)	*Per cent*
Molasses-based	12	35
Grain-based	10	29
Whisky	5	15
Wine-based	3	9
Potato-based	2.5	7
From fruits etc	1.5	5
	34	100

TABLE 4.1.4

Split on areas of the production of potable alcohol

	100% ethanol ($\times 10^6$ hectolitres)	*Per cent*
Asia and Oceania	4.4	13
Africa	0.2	1
Western Europe	17.2	51
Eastern Europe	4.5	13
North America	6.1	18
Latin America	1.6	5
	34	100

References

Aschengreen, N. H. *Process Biochem.* **8,** 23–25 (1969).
Hagen, H. A. *Meeting on Bio-fuels,* Bologna, Italy, June (1981).
Lützen, N. W. *VI int. Fermentation Symp.,* London, Ontario, Canada, July 20–25 (1981).
Nakano, M. *4th int. Fermentation Symp. Kyoto, Japan,* March 19–25 (1972).
Norman, B. E. & Lützen, N. W. *Int. Symp. Cereals,* Carlsberg Research Center, August 11–14 (1981).
Rosén, K. *Process Biochem.* **5,** 25–26 (1978).
Schrøder, P., Novo Industri A/S, DK-2880 Bagsvaerd (personal communication).
Ueda, S. & Koba, Y. *Int. Fermentation Technol.* **3,** 237–242 (1982).

INDUSTRIAL ENZYMOLOGY

The Application of Enzymes in Industry

TONY GODFREY
and
JON REICHELT

M
The Nature Press

Published in the United Kingdom by
MACMILLAN PUBLISHERS LTD (Journals Division), 1983
Distributed by Globe Book Services Ltd
Canada Road, Byfleet, Surrey, KT14 7JL, England

ISBN 0 333 32354 8

Published in the USA and Canada by
THE NATURE PRESS, 1983
15 East 26th Street, New York, NY 10010

Library of Congress Cataloging in Publication Data
Main entry under title:
Industrial enzymology.
Includes index.
1. Enzymes – Industrial applications. I. Godfrey, Tony, 1939– . II. Reichelt, Jon, 1949–
TP248.E51516 1983 660'.63 82–14461
ISBN 0–943818–00–1

Printed in Great Britain

STARCH

J. R. Reichelt

1. Introduction

During the last decade the liquefaction and saccharification of starch-containing raw materials by enzymes have become increasingly more important than traditional acid and acid-enzyme hydrolysis techniques. Enzyme technology applied to the processing of starch provides higher yields, significant improvements in product quality as well as energy savings.

This chapter describes the basic structure and composition of starch and goes on to give a comparison of various native starches, including their gelatinization temperatures. This is followed by a description of the processes of gelatinization, liquefaction, saccharification and isomerization of starches by enzymes. Process parameters are discussed, including indications of substrate and enzyme use levels for the production of maltodextrins, 42–63 dextrose equivalent syrups, high maltose syrups, glucose syrups and isoglucose (high fructose syrups). The formation of amylose–lipid complexes in starches and their effects on gelatinization of liquefaction processes are considered. These complexes are the cause of many problems encountered at the saccharification and final product preparation stages of starch hydrolysate production. Future trends in enzyme starch processing are considered, including the latest developments in the manufacture of isoglucose, together with the recent introduction of a high productivity glucose isomerase. (It should be noted that the objective has been to present the most recent information on the practice of starch conversions).

Starch is the reserve carbohydrate source of plants and is found, for example, in cereal, roots, tubers and palm stem pith. These varied sources yield starches with significantly different chemical and physical properties. Consequently, a number of different techniques are employed in the industrial conversion of these starches to sweeteners.

Starch is found in plant cells as large granules which can be seen under the microscope. These granules are either arranged in concentric layers, as in cereal, or as eccentric layers, as in potatoes. These layers are clearly visible where starch granules have been exposed to heat treatment. Starch is made up of two types of glucose-linked polymers: amylose consists of an unbranched chain of α-1,4 glucoside glucose linked residues, 250–300 units long in

the form of a helix; amylopectin is a branched chain of α-1,4 glucoside and α-1,6 glucoside linked glucose residues up to 1000 units in length. These two polymers are linked together to form a crystalline structure. (*See* Chapter 4.2, Figure 4.2.4.)

In nature amylopectin and amylose are combined in complexes with other cellular components. For example, amylose forms a complex with fatty acids, phospholipids and other substances found in cereal starches. In potato starch amylopectin is joined in a complex with phosphoric acid esters (*See* Figure 4.15.1).

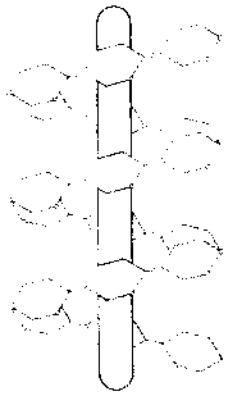


Figure 4.15.1 Schematic representation of an amylose–amylopectin complex.

2. Enzyme processing

The introduction of acid enzyme conversion processes for corn (maize) in the late 1930s provided the technology for the production of noncrystallizing syrups of high sweetness and fermentability for starch processors. In the 1960s amyloglucosidase (glucoamylase) was commercially available and the production of dextrose (D-glucose) followed. During the last decade the liquefaction and saccharification of starch-containing raw materials by enzymes have steadily increased in importance. Because of their greater efficiency and better quality of product, enzyme techniques have largely replaced those of acid hydrolysis.

Enzymes have contributed greatly to the growth of the starch industry by improving the existing processes and also by providing a wide variety of starch hydrolysates with well-defined physical properties and carbohydrate profiles. Glucose isomerase is an excellent example of the application of enzyme technology; its use in the production of isoglucose (high fructose syrups) gives a range of syrups with sweeteners equal to or exceeding that of sucrose. Figure 4.15.2 indicates the basic outline for enzyme starch processing through process sequence and product formation.

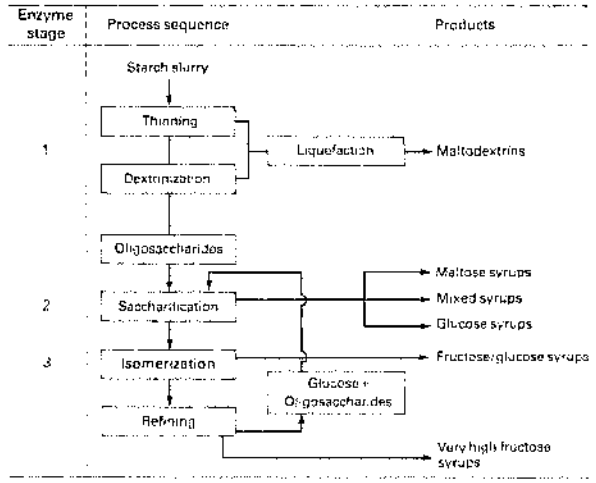


Figure 4.15.2 Main steps in enzyme starch processing.

Table 4.15.1 gives a brief summary of the major starch hydrolysis products and their typical applications.

TABLE 4.15.1
Starch hydrolysis products and applications

Starch product	*Typical applications*
Maltodextrins	Fillers, stabilizers, glues, pastes, thickeners
Mixed syrups (42–63 dextrose equivalent)	Confectionery, soft drinks, brewing and fermentation, jams, conserves and sauces, ice cream, baby foods
High maltose syrups	Hard confectionery
Glucose syrups	Soft drinks, caramel, wine and juice fermentations
Isoglucose (high fructose syrups)	Soft drinks, conserves, sauces, yoghurt, canned fruits

The most important stages in a successful starch conversion are gelatinization and liquefaction. However, prior to processing starch it must first be mixed with water, to form slurries or paste; in industry these should be 25–40 per cent dry substance basis (DSB).

Gelatinization. The starch slurry is heated to above 60°C so that the starch granules swell and burst (Williams, 1968). This process also releases any adhering protein material which then coagulates. The temperature that is required to produce total gelatinization depends mainly on the source of the starch. For example, for waxy corn starch a temperature of 105–110°C will produce complete gelatinization. This gelatinization process produces extremely high

viscosities, and so thinning agents are a necessary addition; they reduce the viscosity and also prevent retrogradation (precipitation) and partial hydrolysis of the starch. The use of cold water to swell the starch granules is not effective enough in distorting the crystalline structure of the granules.

It is essential that the starch is heated above the gelatinization temperature to ensure disruption of the starch granules and to present a suitable substrate for enzyme action. The work carried out by Katz (1928) with X-ray diffraction studies provided the key to the cause of these characteristics. Working with wheat starch, Katz found two stages of gelatinization corresponding to two changes in crystallinity; the initial stage at 60 °C and the second stage between 100–105 °C. Recent work by Kuginiya *et al.* (1980, 1981), Höpcke *et al.* (1980), Eberstein, Konieczny-Janda and Stute (1981), Stute and Woelk (*c*1983) and Konieczny-Janda and Richter (1982) has supported these facts as expressed in their differential scanning calorimetry thermograms and X-ray diffraction measurements on wheat and other native starches. These results are summarized in Table 4.15.2. Differential scanning calorimetry thermograms are measurements of the heat uptake of starch mixtures or slurries during heating.

TABLE 4.15.2
Gelatinization temperatures of different native starches

	Gelatinization temperatures (°C)[1]					
	Method: DSC[2]			*Method: microscope*[3]		
Starch	Onset	Peak	Conclusion	Onset	Peak	Conclusion
Corn starch	65	71	77	65	69	76
Waxy corn starch	65	72	80	64	70	78
Wheat starch	52	59	65	55	61	66
Rye starch	49	54	61	51	54	58
Oat starch	52	58	64	54	58	61
Rice starch	70	76	82	72	75	79
Potato starch	61	65	71	58	64	68
Tapioca starch	63	68	79	64	69	80
Arrow root (Maranta starch)	67	75	85	69	76	84

[1]Höpcke *et al.* (1980).
[2]Differential Scanning Calorimetry DSC-111, Setaram/Lyon (heat uptake during swelling of the starch granules).
[3]Hot stage microscope, Leitz/Wetzlar (loss of birefringence of the starch granules).

Amylose–lipid complexes. During the initial heating of starch slurries there is an endothermic effect between 55–85 °C due to the breakdown of the partial crystalline structure. With non-waxy cereal starches, with normal amylose and lipid content an additional effect between 85–107 °C is apparent. This has been shown to be caused by the dissociation of amylose–lipid complexes, which are naturally present in starch (Konieczny-Janda & Richter, 1982; Kugimiya *et al.* 1980; Konieczny-Janda & Stute, 1981; Stute & Woelk, 1983). The thermostability of these complexes has been shown to increase with increasing chain length and fatty acid saturation number (Konieczny-Janda & Stute, 1981). Recent studies have shown that approximately five to ten per cent lipid material is sufficient to complex amylose almost completely. From this work the complexed part of amylose in native starch granules has been appraised and estimated values of 24 per cent in maize and 33 per cent in wheat starch have been quoted.

For starch processors, the most important factor arising from this work is that the second transition peak measured between 85–107 °C is 'reversible'. This means that after heating the starch paste or slurry to a very high temperature (up to 150 °C) and then subsequently cooling it, the insoluble amylose–lipid complex precipitates again. These complexes are the major cause of hazes and flocculants in saccharification procedures, which are carried out at 60 °C for periods of 48–72 hours. These precipitates are very difficult to remove and usually involve filtration procedures; in some cases they have proved impossible to remove. This is termed 'retrogradation of starch'.

3. Liquefaction

Traditionally the processes of thinning and dextrinization of gelatinized starch were carried out by acid (Palmer, 1981). The starch slurry was acidified to pH 1.5–2 and heated to 140–155 °C for 5–10 minutes. This resulted in complete gelatinization of all starches, and produced hydrolysates which could easily be filtered, however; many reversion products, colour and salts were also produced. This led to the use of acid/enzyme and then enzyme systems for liquefaction, with the use of bacterial α-amylase enzymes from *Bacillus subtilis*. These enzymes are able to operate at temperatures of 85–87 °C and for short periods of time at 90–95 °C.

Initial use of these enzymes presented problems as not all native starches can be gelatinized at temperatures of 90–95 °C, potato and waxy starches being the exceptions. The two stage addition process was introduced to overcome these problems with a pressure or jet cooking stage at temperatures of 140–150 °C for 5 minutes.

Two stage addition batch process. This process involves the addition of bacterial α-amylase (e.g. *Optiamyl*® or *Tenase*® at 0.2–0.4

litres per 1000 kilogrammes starch dry substance basis) to the starch slurry 25–40 per cent dry substance basis to which sufficient calcium ion, 200–400 parts per million (dependent on water hardness), and 300–450 parts per million sodium ion have been added after adjustment to pH 6.8–7.0. The slurry is then heated to 85–90 °C and held for approximately 20 minutes. Pressure cooking at 140 °C for 5 minutes and cooling to 85 °C is followed by the second addition of enzyme (e.g. 1.0–1.8 litres per 1000 kilogrammes dry substance basis *Optiamyl*® or *Tenase*®). This is sometimes called the second liquefaction or dextrinization step. This temperature is held until the desired dextrose equivalent (DE) is achieved. Dextrose equivalent is the term used to characterize the degree of degradation of the starch; it is the reducing power of the starch material as compared with pure dextrose which represents 100 per cent. When the desired dextrose equivalent is reached the starch hydrolysate is heated to 100 °C, and held for 10–15 minutes to ensure that all the enzyme is inactivated.

Similar enzyme processes of this type had widely replaced acid liquefaction methods until the development of heat stable, high temperature α-amylases from *Bacillus licheniformis*.

Continuous enzyme starch liquefaction. The development of thermostable high temperature bacterial amylases from *Bacillus licheniformis* which could operate at sustained temperatures above 95 °C, and withstand temperatures of 105–110 °C for short periods, led to the continuous starch liquefaction process with a single enzyme addition step as shown in Figure 4.15.3.

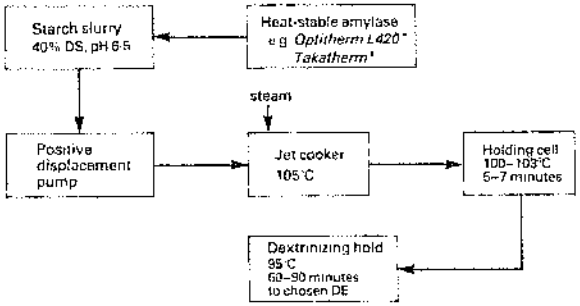


Figure 4.15.3 Continuous enzyme starch liquefaction.

These enzymes have as their cofactor a calcium metalloprotein, in which the calcium is tightly bound to the enzyme. Only low levels of calcium ion are required in processing, 75–100 parts per million, for example. However, allowance should be made for any degree of hardness of the process water. The enzyme is also very low in protease activity, therefore the risk during processing of protein

hydrolysis and subsequent formation of coloured products is reduced. These products increase the number of purification stages and thus the processing costs. Starch substrates with higher protein content can therefore be used for syrup production as a result of the introduction of these thermostable enzymes.

The systems most frequently used have been jet cooker or live steam injection processes operating at 105–110 °C with a single enzyme addition step prior to the jet (typically 1.8 litres per 1000 kilogrammes starch dry substance basis *Optitherm*® or *Takatherm*®, or 0.9 litres per 1000 kilogrammes starch dry substance basis *Optitherm*® *L 420*). This is followed by a short holding time of 5–10 minutes at these temperatures and then flash cooling at 95–100 °C for 1–2 hours until the required final dextrose equivalent has been reached (usually 12–15 dextrose equivalent). Heat treatment at temperatures of 120 °C and above with low pH are required to inactivate the enzyme before saccharification. The jet cooker is usually described as a venturi tube in which the live stean is introduced into the starch slurry with a mixing action that allows instantaneous heating to the required temperature. The length of pipe at the end of the cooker can be varied to achieve short time delays, rather than using holding cells where high temperatures are used (e.g. 140–150 °C).

Although this system has been used successfully for potato and cereal starches, it would not achieve total gelatinization of all other native starches, as previously described above in the section on amylose–lipid complexes; amylose–lipid complexes and their associated problems have been experienced with maize and wheat starches. Miles Kali Chemie, who have overcome these problems, recommend a dual addition process which ensures complete gelatinization and minimizes the retrogradation of starch.

Dual enzyme addition jet cooking process. (*See* Figure 4.15.4.) This process is ideal for the preparation of maltodextrin, 42–63 dextrose equivalent, maltose, high maltose, glucose and isoglucose syrups. Although most of the world's glucose syrups are produced from maize starch, usually by wet milling processes, within the European Community there is a growing use of wheat starch for sweetener production. These starches require much higher temperatures for complete gelatinization than those used in the single addition process.

A starch slurry of 35–40 per cent dry substance basis starch is prepared, and calcium chloride added to achieve 100 parts per million calcium ion in the slurry, with some allowance for the degree of hardness of the process water. The pH is adjusted to 6.0–6.5 and for the first liquefaction step, 0.15–0.3 litres *Optitherm*® *L210* or

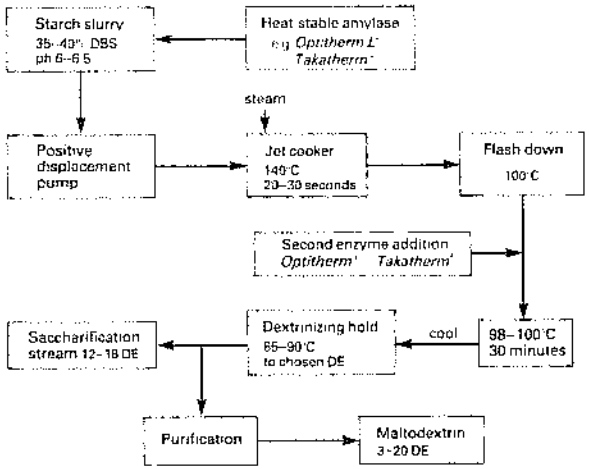


Figure 4.15.4 Dual enzyme addition jet cooking process.

0.075–0.15 litres *Optitherm® L420/Takatherm®* added per 1000 kilogrammes dry substance basis starch. The slurry is then passed through a jet cooker at 140°C and held for a minimum of 20–30 seconds. This time delay can be extended by an additional length of exit tube from the cooker, the time delay being dependent on the process conditions (i.e. rate of throughput). The treated slurry is then flashed down to 100°C and the addition enzyme put in (0.7–1.0 litres *Optitherm® L210* or 0.35–0.5 litres *Optitherm® L420* per 1000 kilogrammes dry substance basis starch). This is held at 98–100°C for 30 minutes. The slurry is then cooled to 85–90°C and held at this temperature until the desired hydrolysate or liquefaction is achieved. This reduction in temperature avoids colour formation and chemical isomerization of maltose to maltulose; also the yield of dextrose following saccharification is higher. Recent work has indicated that the dextrose equivalent level is achieved at a faster rate when the temperature is lowered to 85–90°C and this is currently being investigated.

The slurry is then heated to 120°C and held for 10–15 minutes to inactivate the enzyme; an acid pH 3.8–4.5 will aid this process where applicable. The resulting hydrolysate can be filtered, clarified with carbon, concentrated and dried or spray dried. Where maltodextrin is being prepared of dextrose equivalents of 3–20, or where intermediate dextrose equivalents are required, dextrose equivalents of up to 40 may be achieved using enzyme liquefaction.

The jet process as a continuous process has many advantages for enzyme liquefaction procedures. The process is simple, dependable, offers flexibility, is highly efficient and economical in use. The jet process can increase capacity at minimal capital investment and occupies minimal space. The process improves the control over the liquefaction stage so resulting in improved product uniformity and quality. The resulting low dextrose equivalent hydrolysates contain a minimal consistent level of saccharides which is particularly important where the liquefied starch is to be used for dextrose, high maltose syrup production or other enzymatic conversions.

4. Saccharification

Although the dextrin complex produced from the liquefaction system is commercially valuable for its rheological properties and as a carrier for other food ingredients (*see* Table 4.15.1), it does in fact form the substrate for enzymatic saccharification.

Further hydrolysis of the oligosaccharides is achieved by the use of two saccharifying enzymes: amyloglucosidase (glucoamylase) from *Aspergillus niger* and fungal α-amylase from *Aspergillus oryzae*. These two enzymes used separately or in combination are capable of producing a variety of sweeteners with widely differing sugar profiles (*see* Table 4.15.3).

Amyloglucosidase is an exo-α-amylase and produces glucose from oligosaccharides. It is used for the saccharification of liquefied starch to dextrose (glucose) syrup, and where active on an enzyme liquefied substrate it can produce 96–98 dextrose equivalent syrups.

Fungal α-amylase is an endo-α-amylase and will hydrolyse α-1,4 oligosaccharides to maltose and maltodextriose. It is used where a maltose syrup is required with little dextrose (glucose) production, and is capable of working on low dextrose equivalent substrates. Fungal amylase has a much broader substrate specifity than bacterial α-amylase and is capable of both dextrinizing (liquefying) and saccharifying actions on starch.

Amyloglucosidase and fungal amylase can be used in combination to give high conversion syrups, for example 62–63 dextrose equivalent high conversion syrups with profiles of 30–35 per cent dextrose and 40–45 per cent maltose. The saccharification process should be carried out as soon after liquefaction as practically feasible, and cooled rapidly to the saccharification temperature to avoid retrogradation.

Dextrose (glucose) syrups. (*See* Figure 4.15.5.) For the production of glucose syrups amyloglucosidase enzyme is used for the saccharification of the liquefied starch to dextrose syrup. Refined dextrose syrup with a dextrose equivalent of 97–98 usually has a D-glucose of 95–97 per cent dry substance, with 3–5 per cent higher saccharides, usually maltose or isomaltose.

The final syrup can be spray dried or dried and sold without

TABLE 4.15.3
Summary table for enzyme produced syrups from starch

Process	*Syrup*				
	Glucose	*Maltose*	*High maltose*	*High conversion*	*Isoglucose*
Liquefaction	Thermostable bacterial α-amylase	Thermostable or conventional bacterial α-amylase	Thermostable bacterial α-amylase	Acid/conventional/ thermostable bacterial α-amylase	Thermostable bacterial α-amylase
Saccharification	Amyloglucosidase (AG)	—	—	Amyloglucosidase (AG)	Amyloglucosidase (AG)
	—	Fungal α-amylase	Fungal α-amylase	Fungal α-amylase	—
Isomerization	—	—	—	—	Glucose isomerase
Profile					
Dextrose equivalent	96–98	40–45	48–55	56–68	98
Glucose	95–97	16–20	2–9	22–35	52
Maltose	1–2	41–44	48–55	40–48	—
Fructose	—	—	—	—	42
Isomaltose	0.5–2	—	—	—	—
Maltotriose	—	—	15–16	—	—

Figure 4.15.5 Production of dextrose (glucose) syrups.

further purification. However, the syrup may also be used for the production of pure dextrose by a two or three stage crystallization processes. This syrup is also used as the starting point for the production of isomerized syrups, for example isoglucose (high fructose syrups).

The saccharification process is usually carried out in tanks equipped with agitators because of the long reaction times used in the process (48–96 hours). These tanks are usually used as batch reactors, although they can also be used in series to form tank reactors, but here it is difficult both to control and to obtain very high dextrose equivalent products.

After starch liquefaction the solution should be between 27–40 per cent dry substance. Following a rapid cooling to 60°C the pH is adjusted to pH 3.8–4.5, usually with hydrochloric acid. The amyloglucosidase enzyme is then added while the tank is filling (e.g. 1.0–1.2 litres *Optidex® L150* or 0.75–0.9 litres *Optidex® L200* per 1000 kilogrammes dry substance basis). The temperature must be carefully maintained at 60°C to optimize the reaction rate. Temperatures above 60°C reduce the stability of amyloglucosidase, while reduced temperatures lead to a drop in the reaction rate and so increase the risk of microbial infection. Gentle agitation should be used and after 48–72 hours a final dextrose equivalent of 97–98 should be obtained using this enzyme process.

The reaction should be stopped when the maximum dextrose level is obtained, for if the reaction is continued the glucose level will fall. This is due to the reverse reaction whereby a condensation reaction produces maltose and isomaltose. If the syrup is not ion exchanged following saccharification then for further processing a

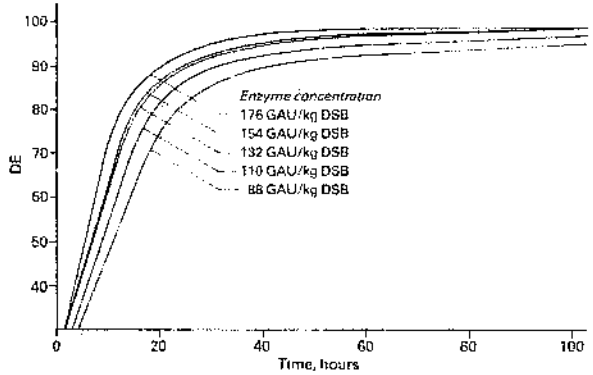


Figure 4.15.6 Relationship between enzyme concentration and dextrose equivalent with time.

heat treatment must be used to inactivate any remaining amyloglucosidase activity. This is achieved by heating to 80 °C for approximately 20 minutes, but pH adjustments will reduce this time. The saccharification time depends on the dosage of enzyme used and Figure 4.15.6 shows the relationship between enzyme concentration and dextrose equivalent with time. Low substrate levels are also converted much more efficiently than high concentration. However, *Optidex*® will convert liquefied starch of up to 50 per cent dry substance basis into glucose. For industrial processes 27–40 per cent dry substance basis is used and normal saccharifications are carried out at 30–35 per cent dry substance basis to achieve high dextrose equivalent syrups at economic costs. Low substrate concentration products would involve higher evaporation costs to achieve the same yields. The recommended pH range of 3.8–4.5 based on experience provides maximum conversion with minimal colour formation, and it also reduces the amount of carbon required in any clarification procedures.

During initial purification the syrup is usually filtered or passed through a separator system to remove insoluble materials, such as fat and denatured protein. The syrup can then be further refined by activated carbon and ion exchange treatments.

Lower dextrose equivalent syrups can be obtained by using amyloglucosidase with the liquefied starch as previously described, and stopping the reaction when the dextrose equivalent reaches approximately 36–42. Such syrups contain glucose, maltose, maltotriose and higher sugars and their typical profile is shown in Table 4.15.3. In the past the majority of these syrups were made by acid/enzyme processes.

Maltose syrups. Prior to liquefaction using heat stable bacterial α-amylase, maltose syrups were produced by saccharification using

malt extract on acid liquefied starch. At present the use of fungal α-amylase provides a process which is more economic than malt extract (cereal β-amylase).

Low glucose-containing maltose syrups. A high maltose, low dextrose syrup can be produced using fungal α-amylase on enzyme liquefied starch suspension and yielding dextrose equivalents of 10–20. The oligosaccharide mixture is concentrated to 38–52 per cent dry substance basis, the pH adjusted to 5.0–5.3 and cooled to 55 °C. Preparation is recommended in a stirred tank system, similar to that for dextrose production. Agitation at slow speed is recommended. Fungal α-amylase is added to the tank as it is filled (e.g. MKC *Fungal Alpha Amylase* at a level of 0.016–0.024 per cent dry substance basis). This level of enzyme addition will give an approximate conversion time of 40–48 hours. Shorter conversion times can be achieved by increasing the quantity of enzyme.

After the desired dextrose equivalent has been obtained, it is essential to inactivate the enzyme to minimize dextrose production. This can be achieved by raising the temperature to 80–85 °C for approximately 20–30 minutes. The final syrup should be processed by conventional plant methods. Finished syrups from this process should have profiles giving dextrose equivalents between 48–52, with 48–52 per cent maltose, and 5–9 per cent dextrose. In some cases maltose levels of 60 per cent have been achieved when using enzyme liquefied starch with fungal amylases.

High conversion syrups. (*See* Figure 4.15.7.) These syrups are produced by the use of fungal α-amylase and amyloglucosidase enzymes. Both these enzymes have similar pH optima and they are able to act simultaneously on starch hydrolysates to produce high

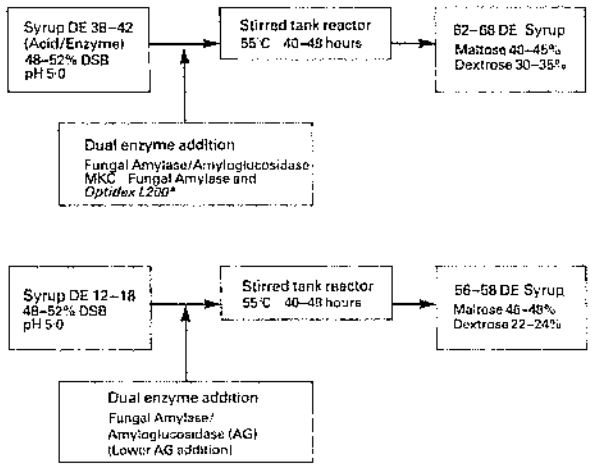


Figure 4.15.7 Production of high conversion syrups.

conversion syrups of 62–63 dextrose equivalents. The syrups are stable enough to resist crystallization at low temperatures and high concentration (80–83 per cent dry substance). With their excellent glucose maltose ratio, colour, flavour, sweetening properties and high fermentability they are widely used in food processing, brewing and the fermentation industries (*see* Tables 4.15.1 and 4.15.3).

In the past these syrups were produced from acid liquefied starch beginning with dextrose equivalents of 38–42. After preparation of a 38–42 dextrose equivalent syrup by acid or enzyme liquefaction the syrup is concentrated to 48–52 per cent (dry substance basis). A lower dextrose equivalent of 14–18 can also be used and these syrups give a higher maltose to dextrose ratio.

The concentrated syrup is adjusted to pH 5.0–5.2 and the temperature lowered to 55°C. Preparation is recommended in stirred tanks, using low speed agitation. The dual enzyme addition (e.g. 6 grammes per 100 kilogrammes dry substance basis MKC *Fungal Amylase-P 40.000* and 700 Glucoamylase Units per 100 kilogrammes dry substance basis *Optidex®-L*) is made as the tank is filled to give an approximate conversion time of 48 hours.

After the desired dextrose equivalent has been obtained the enzymes should be inactivated to minimize dextrose production. This can be achieved by heating to a temperature of 95°C for approximately 10 minutes. Heat exchangers can also be used for this inactivation as they minimize syrup discolouration and increase flexibility during conversion. Contact times of two to three minutes at 100°C result in instantaneous inactivation. Carbon treatment is also very effective at removing fungal α-amylase and amyloglucosidase enzymes for the converted syrup.

However, it is possible by very careful control of enzyme usage, time and plant schedules to produce processed syrups without enzyme inactivation. Table 4.15.4 gives use levels and conversion times required to reach a dextrose equivalent of approximately 62–63 under optimal conditions, using 50 per cent dry substance basis substrate with MKC *Fungal Amylase-P 40.000* and *Optidex-L®*

TABLE 4.15.4
Effect of enzyme concentration on conversion time

Fungal amylase (MKC Fungal Amylase-P 40.000)	*Amyloglucosidase (Optidex-L®)*	*Time (hours)*
9.0 g/100 kg DSB	1050 GAU/100 kg DSB	36–48
6.0 g/100 kg DSB	700 GAU/100 kg DSB	48–60
4.5 g/100 kg DSB	515 GAU/100 kg DSB	60–72

Use levels may vary from plant to plant due to differences in starting substrates and processing conditions.

The syrups are processed by conventional methods and should yield syrups with the following approximate profiles: dextrose equivalent 62–63; dextrose 30–35 per cent; maltose 40–45 per cent; maltotriose 8–10 per cent and the remainder consisting of 20–22 per cent higher saccharides.

5. Isomerization

Isoglucose (high fructose syrups). During the late 1960s in the USA fructose syrups containing 15–42 per cent fructose were produced from starch. These syrups were initially called high fructose corn syrups (HFCS) and more recently high fructose syrups (HFS). Within the European Community these syrups are now called isoglucose. Since 1970 these syrups have been manufactured by enzymatic isomerization of glucose. They have gained an increasing share of the industrial sweetener market, especially in the USA, assisted by the rise in sugar prices in 1974 and 1975. These syrups have approximately the same composition as invert sugar and allow the same sweetness as sugar to be produced from starch sources.

The growth of fructose syrups in the USA has not been paralleled within the European Community. Production of isoglucose was severly limited by the European Commission's quota system and then levies on starch sources, while the Common Agricultural Policy will continue to restrict product development. However, second generation fructose syrups are now being produced with 55–60 per cent fructose in the USA and Japan, and they will

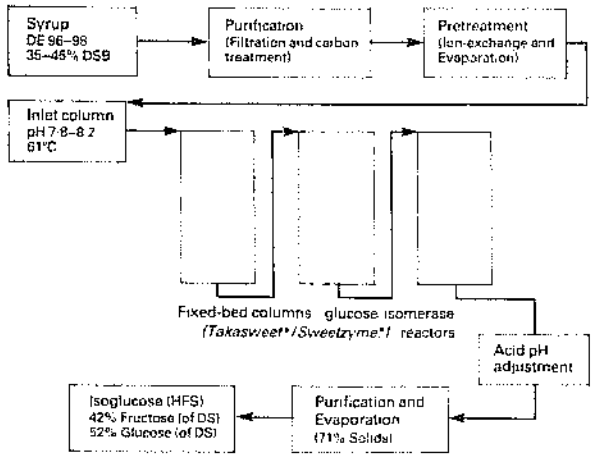


Figure 4.15.8 Typical process layout for isoglucose production.

probably soon be exported as production increases.

Glucose isomerase. Glucose isomerase enzymes catalyse the isomerization of D-glucose to D-fructose, a reaction that is reversible. Fructose formation is favoured by using alkaline pH conditions and at equilibrium, ratios or 52 per cent fructose, 48 per cent glucose are achieved. The enzyme is a thermophillic metalloenzyme and requires traces of magnesium as cofactor. The isomerization process is summarized as the transfer of two electrons from one carbon to the next, and the conversion of aldose to ketose. (For further information in this area *see* Chapter 2, 'Kinetics', pages 35–37).

In nature there is no reason why there should be an enzyme to convert glucose into fructose. To achieve this goal one would have to develop an enzyme specifically for the purpose, which is not yet possible, or use what is available – xylose isomerase. Glucose isomerases should really be categorized as D-xylose isomerases which have D-glucose-isomerase as a side or secondary activity.

This means that with batch reactors enzyme dosages are required using soluble enzymes for the conversion, resulting in a costly process. Glucose isomerases are intracellular enzymes and therefore do not yield the same quantities of product as extracellularly produced enzymes, and therefore have higher production costs. In order to overcome the inefficiency of the enzyme and still run a commercially viable process it was essential to keep reusing the enzyme. This repeated usage was achieved by enzyme immobilization. Now, several of the large enzyme manufacturers produce immobilized glucose isomerases. There are excellent account of these earlier developments in this process technology including mathematical models, process criteria and plant design (Seidman, 1977; Antrim *et al.*, 1979; Hemmingsen, 1979).

Most of the glucose isomerase enzymes used today have been specially developed for use in continous fixed bed column processes with downward substrate flow through the columns. In addition to a continous flow the column process gives short syrup-enxyme contact and allows optimal conditions to be selected for enzyme productivity without significant byproduct formation. This gives a clear colourless syrup and reduces purification costs for the final syrup. Contact time with the enzyme determines the amount of fructose produced and the whole system requires very careful control.

Process parameters. The process parameters which affect activity, stability and productivity have been summarized in Chapter 2, 'Kinetics', Figure 2.61. This figure highlights both the complexity of the system and the need for sophisticated process controls and

analytical techniques to ensure adequate substrate quality, and to maintain pH, temperature and product quality.

The activity of the enzyme is usually expressed as Immobilized Glucose Isomerase Column Units (IGICU). For design purposes the reactor columns are assumed to be plugged flow reactors and calculations based on mathematical models have been drawn up to evaluate enzyme performance, using bed height, particle size, pressure drop, pH and temperature. The activity determined experimentally under defined conditions for the substrate is the initial rate of reaction (i.e. the quantity of fructose formed per unit time per weight of enzyme) starting with a fructose-free substrate. Stability represents the amount of activity retained over time. Productivity (total product produced per quantity of enzyme in a given time period) is a result of the combined effects of activity and stability. In industry productivity is usually defined as the kilogrammes of fructose produced per kilogramme enzyme during its lifetime. The productivity of most first generation glucose isomerases is quoted as between 2000–4000 kilogrammes per kilogramme enzyme.

The effects of temperature on activity, stability and product formation are shown in Table 4.15.5. A temperature of approximately 60°C usually ensures adequate activity and stability while reducing the risk of microbial infection and at an economic level. Figure 4.15.11 shows how a change in temperature or ±1°C from the optimum will affect column productivity significantly.

TABLE 4.15.5
Effects of temperature on activity and product formation

Temperature	t_l *hours*	*Design enzyme lifetime*	*Productivity kg DS/ kg enzyme (200 IGICU/g)*	*Total enzyme Bed volume for 100 tpd-plant*
65°C	350	$2 \times t_l$	1130	9.2 m^3
		$3 \times t_l$	1300	11,7 m^3
61°C	800	$2 \times t_l$	1820	12.6 m^3
		$3 \times t_l$	2100	16.1 m^3
60°C	1000	$2 \times t_l$	2090	13.6 m^3
		$3 \times t_l$	2430	17.4 m^3
57.5°C*	1800	$2 \times t_l$	3100	16.2 m^3
		$3 \times t_l$	3600	20.8 m^3

*This temperature is below the 60°C, recommended and is included here only for illustration purposes. (Data from Novo *Sweetzyme Q*®).

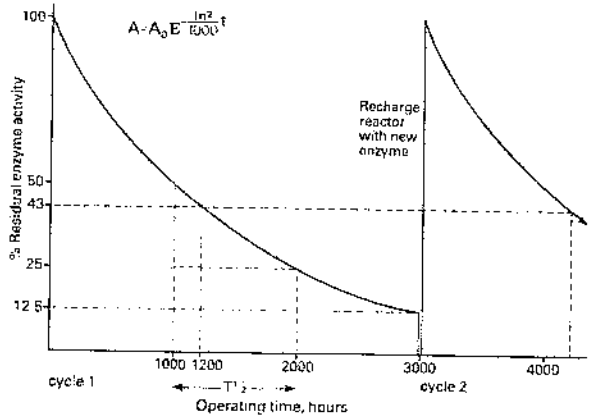


Figure 4.15.9 Activity decay of immobilized glucose isomerase.

pH control is very important and it is recommended that all pH readings are measured at 25 °C. Unfortunately the optimum activity of glucose isomerase and the catalytic stability do not share the same pH value. With soluble enzymes this is not a problem, but in the case of immobilized enzymes in a fixed bed column, the pH of the substrate must be maintained at a constant value to ensure adequate column productivity. In industry this is normally at approximately three half lives (*see* Figure 4.15.9).

The concentration of the substrate syrup must also be controlled. Due to the high solids content there is a tendency of the enzyme particles to show an increased diffusion resistance and this must be reduced. Conversely, low substrate levels also lower activity and increase the risk of microbial fouling in the columns. Dry substance contents of the substrate syrup should be in the order of 38–45 per cent with an optimum level around 38–40 per cent.

The feed syrup should be processed and purified by filtration, separation, carbon treatment and an ion/cation exchange. It is then concentrated by evaporation, which also reduces both the oxygen content and byproduct formation causes of loss of activity of the column. The highest possible dextrose equivalent should be used in the syrup, normally 96–98, for maximum conversion. Reactors run in series or in parallel provide a smooth productivity flow and facilitate process control during column regeneration. The columns are staggered so that they do not all require regeneration at the same time.

Plant design is based on the proposed throughput and this will determine the size of plant the number of enzyme reactors which will give this level of productivity. To achieve an economic operation a number of factors must be taken into account and these

include enzyme activity losses, pressure drop over the enzyme bed, syrup residence time, flow distribution conversion rates and column regeneration cycles with process control. Plant design is a very skilled process and it is advisable to contact the respective enzyme manufacturers for specific advice on plant layout and reactor design recommendations.

Production process. The saccharified feed syrup should have a dextrose equivalent of 96–98, and a dextrose content of 94–96 per cent, 93 per cent being considered the absolute minimum. The syrup is refined by filtration and activated carbon treatment, then passed through ion exchange resins. The refined syrup should be free from heavy metal ions, with calcium less than one part per million. It is then concentrated to 40 per cent dry substance bases. The pH adjusted to that recommended (e.g. pH 7.8 for *TakaSweet®* or 8.2 for *Sweetzyme Q®*) and the temperature brought up to 61 °C just before the syrup enters the isomerization columns. The temperature will drop slightly as the syrup passes through the columns leaving the column at a temperature of around 59 °C. An example of a typical process layout is shown in Figure 4.15.8.

During isomerization byproduct formation is a function of temperature, pH and time. Under the reaction conditions described, residence times of 0.8–4 hours provide the recommended minimum and maximum. Any colour generated during the process can be removed by activated carbon treatment during final product purification.

The activity of the immobilized enzyme decreases with time and it is necessary to control the flow rate during the run in order to achieve the correct degree of conversion which is usually 42–45 per

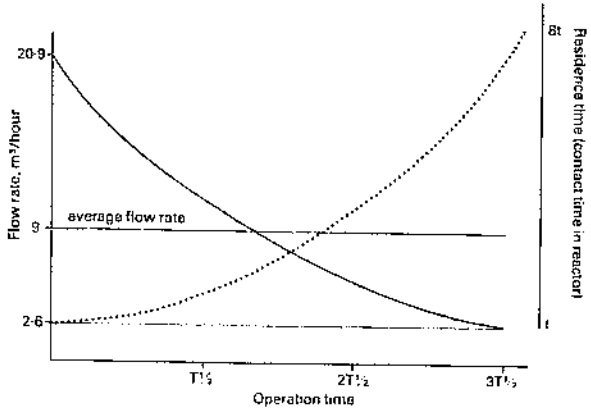


Figure 4.15.10 Flow variation in single column operation.

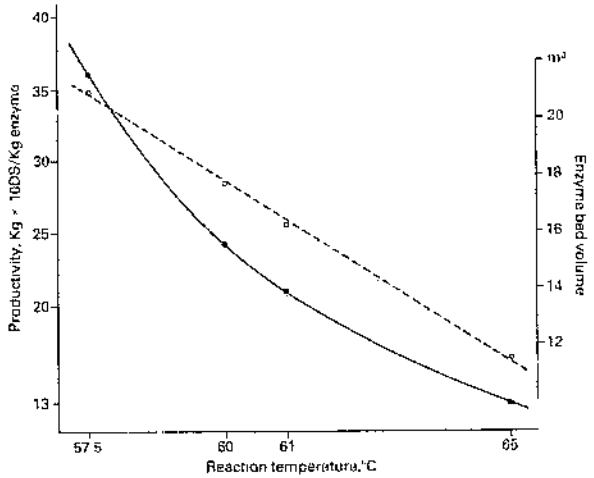


Figure 4.15.11 Effect of temperature on half life productivity.

cent. For production purposes it is essential to know the activity decay curve. This is shown in Figure 4.15.9.

Variation of flow rate (and column residence time) with operation time for a single column is shown in Figure 4.15.10. Figure 4.15.11 shows the effect of temperature on half life productivity and reactor size.

After isomerization, the syrup is blended and the pH adjusted to pH 4.0–5.0. The syrup is then passed through carbon and ion exchange treatments before evaporation to approximately 70–71 per cent dry substance. The syrup is usually sold in bulk and held at 30°C during storage. A typical fructose syrup will have the following approximate profile before ion exchange treatment: fructose 42 per cent dry substance, glucose 52 per cent dry substance, 6 per cent other saccharides.

This process of the isomerization of glucose to fructose by glucose isomerase represents the pinnacle of enzyme technology. The precise control of enzyme activity (productivity) and process control has been achieved by using the latest analytical techniques such as high pressure liquid chromotography (HPLC) and microprocessor control systems. These developments have enabled a low efficiency process to become an acceptable commercial operation.

Editors' note: The detailed operating formulae for calculating analytical and operating criteria, enzyme bed volume, the average activity and productivity of glucose isomerase, are specific to each source of this enzyme and should be obtained direct from the manufacturing supplier. Please refer to Data Index 2.

6. Recent developments

Maltodextrins have recently been defined in the UK by the Ministry of Agriculture, Fisheries and Food (MAFF) through their Food Additives and Contaminants Committee (FACC). The Committee has issued the 'Report on Modified Starches' (FAC/REP/31) in which maltodextrins are defined as starch hydrolysates with dextrose equivalents of 3–20. Maltodextrins are to be excluded from these regulations, but may be only obtained from starch by enzymatic and/or acid hydrolysis, to yield purified aqueous solutions to nutritive saccharides or the subsequent dried product. The enzymes used for the production of maltodextrins, bacterial α-amylasas, are also under review by the Committee, and a summary table of the Committee's recommendations is given in Chapter 3.1.

Miles Kalie-Chemi AG have recently introduced a second generation immobilized glucose isomerase enzyme, *Optisweet® 22*, with a standard productivity of 22,000 kilogrammes dry substance per kilogramme of enzyme. The enzyme has excellent flow characteristics, with smaller particle size and higher space velocities than conventional glucose isomerase enzymes which result in lower fixed bed heights. Current reactors could be used without expensive modifications to increase plant capacity, beyond that of conventional glucose isomerase. New plant designs, however, could use much smaller reactors with reduced capital costs and lower spatial requirements. (A 0.7 metre bed height of *Optisweet® 22* is equivalent to using conventional glucose isomerase with a 3.5–3.7 metre bed height on start-up, with conventional glucose isomerase at pH 7.5 and a substrate inlet temperature of 60°C in a 2-bar test pressure reactor.)

7. Future trends

With the development of a second generation of high productivity immobilized glucose isomerase and the continuous liquefaction processes developed using high temperature heat stable bacterial α-amylases, an immobilized high productivity amyloglucosidase enzyme for saccharification is required for a total continuous system. This enzyme would need to match the current conversion levels attained in batch saccharification using soluble enzyme for high dextrose equivalent syrups to be produced at an economic level.

Several major enzyme manufacturers are working in this field and these developments cannot be too far away. A combined continuous starch process would give starch processors flexibility. With the advances in microprocessor and analytical control systems it would also offer much greater process control.

The development of debranching enzymes such as pullulanase from bacterial sources which have a higher temperature optimum than those currently available could lead to increased dextrose–glucose yields. This occurs during saccharification breaking down any remaining 1,6 linkages, when used with amyloglucosidase.

References

Antrim, R. L., Colilla, W. & Schnyder, B. J. in *Applied Biochemistry and Bioengineering*, Vol. 2 *Enzyme Technology* (ed. Lemuel, B., Wyngard, J., Katchalski Katzik, E. & Goldstein, L.) (Academic, New York, 1979).

Fullbrook, P. D. in *Nutritive Sweeteners* (ed. Birch, G. G. & Parker, K. J.) (Applied Science, London, 1981).

Hemmingsen, S. H. in *Applied Biochemistry and Bioengineering*, Vol. 2 *Enzyme Technology* (ed. Lemuel, B., Wyngard, J., Katchalski Katzik, E. & Goldstein, L.) (Academic, New York, 1979).

Hollo, J. & Szejtli, J. in *Starch and its Derivatives* (ed. Radley, J. A.) (Chapman & Hall, London, 1968).

Höpcke, R., Eberstein, K., Konieczny-Janda, G. & Stute, R. *Stärke* **32,** 397ff (1980).

Howling, D. in *Sugar, Science and Technology* (ed. Birch, G. G. & Parker, K. J.) (Applied Science, London, 1979).

Katz, J. R. in *A Comprehensive Survey of Starch Chemistry* (ed. Walton, R. P.) (Chemical Catalogue Co., New York, 1928).

Konieczny-Janda, G. & Stute, R. *Differential Scanning Calorimetry of Starches, Part 2 Investigation on Starch-Lipid-Complexes* (Paper presented at the 32nd Starch Convention, Detmold, 1981).

Konieczny-Janda, G. & Richter, G. *Enzymatic Degradation of Starch-Lipid-Complexes during Liquefaction of Starch and Synthetic Amylase-Lipid-Complexes*, (Symp. Use of Enzymes in the Food Industry, Paris, 1982).

Kugimiya, M. & Donovan, J. W. *J. of Fd Sci.* **46,** 765 (1981).

Madsin, G. B. & Norman, F. E. in *Molecular Structure and Function of Food Carbohydrates*, 50 (ed. Birch, G. G. & Green, L. F.) (Applied Science, London, 1973).

Palmer, T. P. in *Nutritive Sweeteners* (ed. Birch, G. G. & Parker, K. J.) (Applied Science, London, 1981).

Seidman, M. in *Developments in Food Carbohydrate 1*, 19–42 (ed. Birch, G. G. & Schallenberger, R. S.) (Applied Science, London, 1977).

Stute, R. & Woelk, H. U. German Pat. DE 2231 (*c*1983).

Takasaki, Y. & Yamanobe, T. in *Enzymes and Food Processing*, 73 (eds Birch, G. G. Blackebrough, N. & Parker, K. J.) (Applied Science, London, 1981).

Williams, J. M. in *Starch and its Derivatives* 4th edn (ed. Radley, J. A.) (Chapman & Hall, London, 1968).

Enzymes in Food Technology

Second edition

Edited by

Robert J. Whitehurst

and

Maarten van Oort

WILEY-BLACKWELL

A John Wiley & Sons, Ltd., Publication

This edition first published 2010

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing programme has been merged with Wiley's global Scientific, Technical, and Medical business to form Wiley-Blackwell.

Registered office
John Wiley & Sons, Ltd., The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Editorial office
2121 State Avenue, Ames, Iowa 50014-8300, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data

Enzymes in food technology / edited by Robert J. Whitehurst and Maarten van Oort. – 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-4051-8366-6 (hardback : alk. paper)
1. Enzymes–Industrial applications. 2. Food industry and trade. I. Whitehurst, Robert J.
II. Oort, Maarten van.
TP456.E58E63 2010
664'.024–dc22

2009011853

A catalogue record for this book is available from the British Library.

Set in 10/12 pt Times by Aptara® Inc., New Delhi, India
Printed in Singapore by Markono Print Media Pte Ltd

1 2010

6 Enzymes in bread making

Maarten van Oort

6.1 INTRODUCTION

Baking is a common name for the production of baked goods, such as bread, cake, pastries, biscuits, crackers, cookies, pies and tortillas, where wheat flour is both the most essential ingredient and key source of enzyme substrates for the product. Wheat's popularity is also due to the ease at which it grows under a variety of moderate temperature conditions, soils, regions and seasons.

All wheats belong to the genus Triticum, of the Gramineae family. Common wheat (Triticum aestivum) and durum wheat (Triticum durum) are the two major wheat groups which are used for food applications. Bread is the product of baking a mixture of flour, water, salt, yeast and other ingredients. The basic process involves mixing of ingredients until the flour is converted into a dough, followed by baking the dough into bread.

The aims of the bread-making processes are to produce dough that will rise easily and have properties required to make good bread for the consumer. To make good bread, dough made by any process must be extensible enough to expand during proofing. Bread dough must also be elastic. Elastic dough has the strength to hold the gases produced while rising and stable enough to hold its shape and cell structure.

6.1.1 Wheat

Wheat is one of the major crops in the world with an annual production of around 550 million ton. Depending on the region, either corn, wheat or rice is the dominant food grain.

Compared with corn and rice, wheat contains the unique gluten-forming proteins, which are intrinsically linked to baking.

6.1.2 Wheat flour constituents

Endosperm is the interior of a wheat kernel and makes up about 83% of the whole grain of wheat and once ground down to a powder, is flour.

Flour contains mostly starch. However, other components in the flour also clearly affect its properties. The main components are starch (70–75%), protein (9–14%), lipids (1–3%), non-starch carbohydrates (1–2%), ash (around 0.5%), lipids (1–2.5%) and moisture (13–14%).

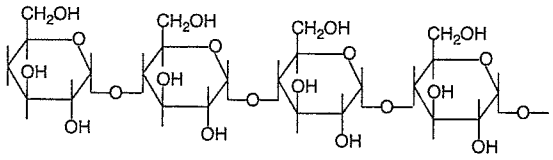


Fig. 6.1 Schematic representation of amylose.

6.1.3 Starch

Starch is the most abundant component and the most relevant reserve constituent of cereals. Starch consists primarily of D-glucopyranose polymers linked together by α-1,4 and α-1,6 glycosidic bonds. The glycosidic linkages in starch are in the alpha (α) configuration. This is determined by the orientation of the hydroxyl (—OH) group on C1 of the pyranose ring. The α-linkage allows starch polymers to form helical structures. The significance of the helical geometry becomes clear when starch is compared with cellulose. Cellulose is a glucose polymer with β-1,4 linkages between the glucopyranose subunits. Because of this configuration, cellulose forms sheeted structures whereas starch polymers are usually helical. This has a strong effect on physicochemical properties and enzyme susceptibility.

In starch there are two types of polymers, amylose and amylopectin.

Amylose (Fig. 6.1) is essentially a linear polymer, although it is known that some branches are present.[1] From the shape of the amylose chain, it becomes clear that this molecule forms helices due to the α-linkages.

Amylopectin (Figs 6.2(a) and (b)) is a branched molecule and is on average a much larger polymer than the amylose polymers.

The structural differences between the two polymers determine the differences in starch properties and starch functionality. Some of the functional differences are listed in Table 6.1.

Although amylose is usually illustrated as a linear chain, the molecule is often helical. The interior of the helix contains hydrogen atoms and can be considered as hydrophobic. This in turn allows amylose to form complexes with free fatty acids, fatty acid components of certain lipids, some alcohols and also iodine.

Complexation with lipids is an especially well-known property of the amylose helix and can alter the properties of starch, such as gelatinization temperature, viscosity and retrogradation, considerably.

Native starch granules are insoluble in cold water. However, upon heated with water, granules undergo a substantial change resulting in a complete change of properties and behaviour. This irreversible process is named gelatinization, which in wheat occurs between 52°C and 85°C.

During heating in water, linear amylose molecules start leaking out of the starch granule, and with continued heating additional amylose and also amylopectin leach out.

After heating, solubilized starch polymers and remaining insoluble granular fragments start to reassemble into an organized structure again. Ultimately, a crystalline structure is formed. This process is known as retrogradation. Linear amylose molecules have a greater tendency to reassociate and form crystalline and gel structures than the larger amylopectin molecules.[2]

The process of retrogradation is closely linked to the problem of staling in baked products. This will be discussed further in Section 6.2.

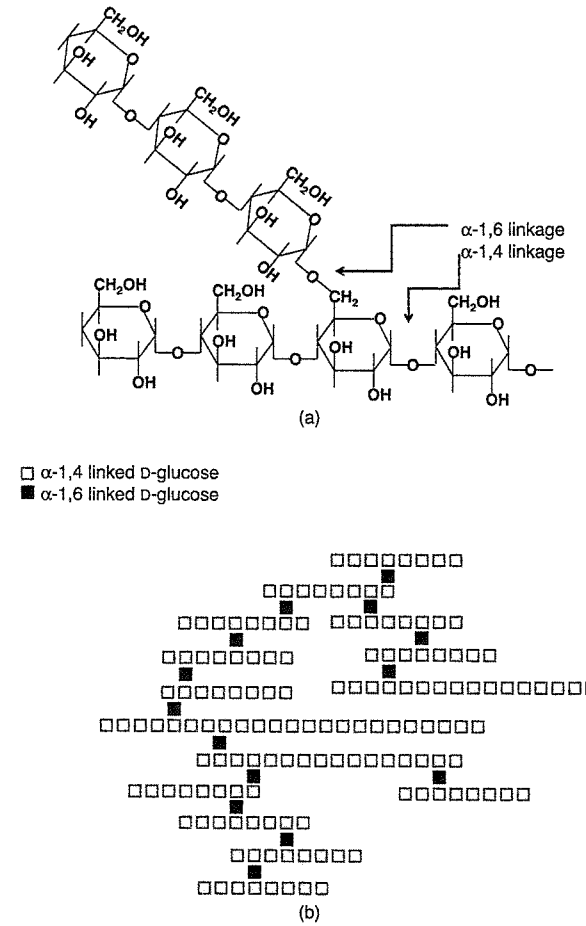


Fig. 6.2 Schematic representations of amylopectin, showing the 1,4 and 1,6 glucosidic linkages (a) and what that means for the whole structure (b).

Table 6.1 Different characteristics of amylose and amylopectin.

Characteristic	Amylose	Amylopectin
Shape	Essentially linear	Branched
Linkage	α-1,4 (±1/1000 α-1,6)	α-1,4 and α-1,6 (±1/25)
Molecular weight	Typically 10^5–10^6	10^7–10^9
Films	Strong	Weak
Gel formation	Firm	Non-gelling/soft
Colour with iodine	Blue	Reddish brown

6.1.4 Gluten

Wheat protein is seen as the most important factor governing bread-making quality.[3–5] A high protein content is related to good bread-making quality. There are also a number of other factors, together described as protein 'quality' that determine the bread-making potential of a wheat flour.[6]

In wheat kernels, 80% of the protein is found in the endosperm. Wheat flour proteins have been classified into four types based on their solubility[7]: albumins, soluble in water; globulins, soluble in salt solutions; gliadins, soluble in 70% ethanol; and glutenins (partly) soluble in dilute acid or alkali. The bulk of the protein is formed by the latter two: gliadins and glutenins. All these fractionations lead to broad overlapping classes of proteins, indicating the complexity and variability of the various classes. There is also only a limited relationship between the amounts of each of the Osborne fractions and the final bread-making quality.[8]

When flour is mixed with water, a viscoelastic mass is formed. From this mass, starch can be washed out and then gluten remains. On a dry basis, gluten contains around 70–85% protein, 5–15% carbohydrates (starch and non-starch polysaccharides (NSP)), 3–10% lipids and 1–2% ash.[5,9]

Gluten proteins contain relatively high amounts of glutamic acid, which exists as glutamine in the wheat kernel, proline, hydrophobic amino acids and the sulphur-containing cysteine. Glutamine and also the hydrophobic amino acids ensure sufficient hydrogen bonding during mixing. This helps in film formation. Proline, because of its ring structure, disrupts the α-helix formation that normally occurs in protein polymers. The result is a higher percentage of β-sheet structure, which helps in giving gluten the necessary elasticity. The sulphur-containing amino acids ensure the formation of disulphide bridges between protein chains and also within protein chains, giving the protein network sufficient strength.

The effect of all these amino acids together is a rapid continuous film formation which gives wheat dough its unique viscoelastic properties and gas holding capacity.[10]

The viscoelastic behaviour of wheat gluten sets it apart from other grains or other vegetable protein sources.[11] The formation of gluten is the result of interaction between the two major classes of wheat proteins, the gliadins and glutenins. When flour is mixed with water, a viscoelastic mass is formed, to a large extent due to this interaction.

The sulphur-rich glutenins are able to form polymeric networks. The sulphur-poor gliadins are present mainly as monomers. Glutenins consist of a high molecular weight (HMW) group and a low molecular weight (LMW) group of proteins. Both groups are able to form disulphide polymeric networks. Gliadins behave as a viscous liquid, whereas glutenins behave as a cohesive elastic solid.[12] In the gluten network, these properties are combined into a protein network with viscoelastic properties, which enables gas cells to be retained in a dough during the bread-making process. This highly specific property allows wheat flour to be used for production of a variety of yeast leavened or chemically leavened products.

The composition and strength (quality) of the gluten is partly determined by the presence and abundance of individual glutenin subunits. The relative amounts of each subunit are determined by genetic factors, growing conditions and fertilization. Glutenin polymers are very heterogeneous in composition and size. Certain fractions of the HMW glutenins are highly correlated with loaf volume, whereas others are not.[13,14] A glutenin fraction that is insoluble in sodium dodecyl sulphate (SDS) solution, has been called glutenin macropolymer (GMP).[15] This fraction has been studied in great detail and evidence has been presented showing that this GMP can be seen as a wheat quality parameter.[16]

A large number of HMW glutenin subunits have been sequenced, relations between various proteins and protein fractions have been established; the influence of glutenin subunits and combinations of subunits on bread-making quality has been investigated.[17] Nevertheless, there is still a great deal unknown, mainly due to the complexity of the gluten network and the large amount of various subunits in the network.

Oxidizing and reducing agents have shown to affect the gluten structure. Ascorbic acid, potassium bromate and cysteine have been used for decades as dough improvers. Modification of the redox status of disulphide bonds or thiol groups significantly changes the polymerization of glutenin subunits of wheat gluten. This in turn affects the mechanical and viscoelastic properties of the dough.

6.1.5 Non-starch polysaccharides

NSP originate from the cell wall of the aleurone and the endosperm of the wheat kernel. NSP represents different polysaccharides. One group of polysaccharides is built up by pentose sugars and some hexose sugars. These are called pentosans and main components are pentose sugars arabinose (A) and xylose (X). The structure of arabinoxylans (AX) is a linear backbone of β-1,4 linked xylose residues carrying single arabinose residues on their C-3 position or both on C-2 and C-3 positions. The molecular weight of AX ranges from 20 000 to 5 000 000 D.[18,19] AX together with cellulose, β-glucans, arabinogalactan-peptide and other minor constituents like galactomannan, glucomannan and xyloglucan are referred to as NSP. Some arabinose residues are esterified with ferulic acid (FA). FA (4-hydroxy-3-methoxycinnamic acid) is the natural component of water extractable AX (WE-AX) and water unextractable AX (WU-AX), although free, soluble bound FA and insoluble bound FA all have been found in flour and in gluten.[20]

Despite their low content in flour (2–3% at normal extraction rates), pentosans are very important in determining dough properties, gluten quality and final bread quality.[21,22] Due to their high water binding capacity, AX play a regulatory role with respect to the water economy in bread making.[23] Furthermore, gluten properties and also dough properties are influenced by interactions between proteins and pentosans.[24]

WE-AX have some unique physical properties, such as binding up to 10 times its own weight of water,[18,25] formation of highly viscous solutions[26] and gels due to covalent cross-linking.[27,28] All these properties have direct functional implications in gluten formation and dough properties. In general, it is believed that WE-AX have a positive effect on bread making,[29] whereas WU-AX are reported to have a strong negative effect on bread-making quality.[30–32] However, Wang[24] has shown that both WE-AX and WU-AX have similar effects on gluten yield, GMP yield, on the composition of gluten and GMP, and on the properties of gluten and GMP.

The negative effects of WU-AX (and WE-AX) on bread-making quality is the main reason why arabinoxylan-modifying enzymes are successfully implemented in almost all bread-making procedures all over the world.[33]

Thanks to the feruloyl adducts (Fig. 6.3), pentosans are subject to oxidative gelation.[26,27] Many studies have been carried out on the oxidative gelation of both WE-AX and also WU-AX.[22,28,34–36] Among various oxidative systems, peroxidases and laccases have been successfully applied as gelling agents of AX solutions.[28,37,38]

This oxidation is generally assumed to occur through cross-linking of free FA residues.

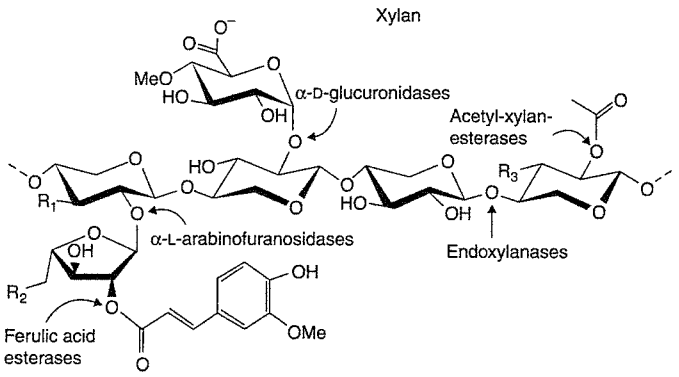


Fig. 6.3 Plant xylan structure with bound ferulic acid.[26]

One possible mechanism is through dimerization of FA residues on adjacent arabinoxylan chains. Proteins also participate in the gelation process, since the gel fraction contains about 25% protein and is dissolved by proteases.[27,35] This indicates that FA is also coupled to the protein, most likely to tyrosine and/or cysteine residues. In total three different possible mechanisms were suggested as explanation for this phenomenon[35,39]: Cross-linking of two FA moieties through their aromatic rings; cross-linking of a FA reside with the ring structure of tyrosine amino acid moieties; or cross-linking of two tyrosine rings.

In feruloylated AX, laccase catalyzes gelation by dimerization of feruloyl esters. Although there was no evidence for a coupling of thiol compounds to phenoxyl radicals, it has been proposed that these phenoxyl radicals, oxidized by laccase from that FA, are involved in the conversion of the thiol compounds into disulphide bonds and thus influence the overall dough firming which is seen upon use of laccase.[36,40]

6.1.6 Lipids

Wheat flour lipids form a highly heterogeneous group of molecules with different chemical structures and compositions. The lipids can be divided into free lipids and bound lipids.[41] Both fractions contain polar and non-polar components. Polar lipids can be divided into glycolipids and phospholipids. The glycolipids are predominantly composed of monogalactosyl diglycerides (MGDG) and digalactosyl diglycerides (DGDG), whereas lysophosphatidylcholines (LPC) and phosphatidylcholines (PC) are major components of the phospholipid class of lipids. Both classes are also highly heterogeneous with respect to the position and the structure of the esterified fatty acids. The non-polar lipid class is mainly composed of triglycerides.[42] The fatty acids in the various lipids are dominated by linoleic acid, whereas other fatty acids, such as palmitic and oleic acid are found in lower levels.[43]

Bound lipids are mainly bound to starch and to a lesser extent to proteins. Starch lipids (around one third of total) comprise mostly LPC. These lipids form inclusion complexes with amylose during gelatinization, whereby the LPC fatty acid chains form complexes with the hydrophobic interior of the amylose α-helices.[44] Such complexes may already exist in the native starch as well. These lipids are therefore effectively unavailable before most of the starch is gelatinized.

The non-starch lipids consist of similar levels of polar and non-polar lipids. Part of this fraction is most likely bound to proteins.[45]

Polar lipids are known to play a role in dough stability and processing tolerance of yeast-leavened baked products. The ability of the polar flour lipids to form lipid monolayers at the gas/liquid interface is believed to positively influence the gas retention in the dough. Furthermore, polar flour lipids and gluten proteins interact as well. This interaction is also believed to positively affect gas retention.[46,47]

6.2 ENZYMES IN BREAD MAKING

Bakery products have undergone radical improvements in quality over the past 10 years in terms of flavour, texture and shelf life. The usage of enzymes is the biggest contributor to these improvements. Among the enzymes used in food applications, those used in bakery industry constitute nearly one-third of the market.

Baking enzymes are used as flour additives and they are used in dough conditioners to replace chemical ingredients and to perform other functions in a label-friendly way.

The baking industry predominantly makes use of five types of enzymes (Table 6.2). Amylases are used to convert starch to sugar and to produce dextrins. For strengthening and bleaching of the dough, oxidases are used. Hemicellulases and proteases are the enzymes which have an effect on wheat gluten. While hemicellulases improve gluten strength, proteases reduce gluten elasticity. All these enzymes together play an important role in maintaining bread volume, crumb softness, crust crispiness, crust colouring or browning and in maintaining freshness.

6.2.1 Amylases

α-amylases (**EC 3.2.1.1**) are the enzymes which are most frequently used in bakeries.[48] The reasons for this are their positive influence on bread volume, crumb grain, crust and crumb colour, flavour development and anti-staling effect.[49,50] There is also evidence that amylases have an effect on dough development.[51]

6.2.2 Classification

Amylases belong to the family of glycohydrolases (GH), based on structural and amino acid similarities.[52] Various amylases are found in GH families 13, 14 and 15. Besides α-amylases

Table 6.2 General effect of enzymes in bread making.

	Improved gluten network	Gas retention/ increased volume	Improved colour and flavour	Improved crumb structure	Improved shelf life properties
Amylase		X	X	X	X
Protease	X				
Xylanase	X	X		X	
Oxidase	X	X		X	
Lipase	X	X	X	X	

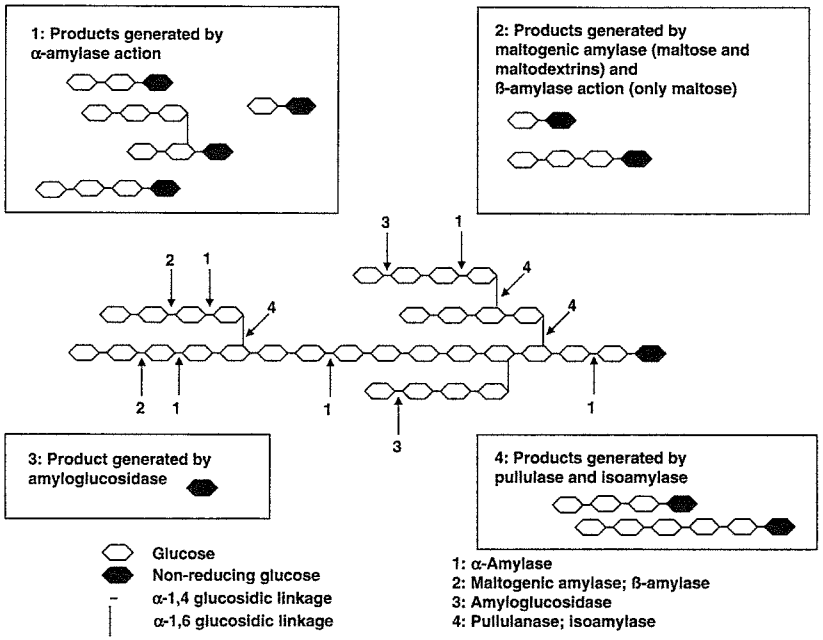


Fig. 6.4 Attack sites and breakdown products for various starch degrading enzymes.

(1,4-α-D-glucan glucanohydrolase **EC 3.2.1.1**), also maltogenic amylase (**EC 3.2.1.33**), β-amylase (**EC 3.2.1.2**), amyloglucosidase (also named glucoamylase, **EC 3.2.1.3**), pullulanase (**EC 3.2.1.41**) and isoamylase (**EC 3.2.1.68**) belong to these families. In the above schematic overview (Fig. 6.4) the action of these enzymes on a starch molecule is shown.

α-Amylase is an endo-acting enzyme that randomly hydrolyzes the α-1,4 glucosidic linkages in polysaccharides, resulting in short chain dextrins. The α-amylases degrade damaged starch in wheat flour into small dextrins of DP2-DP12,[9,53] thus allowing yeast to work continuously during dough fermentation, proofing and the early stage of baking. This results in improved bread volume and crumb texture. In addition, the small oligosaccharides and sugars such as glucose and maltose produced by amylases enhance the reactions for the browning of the crust and baked flavour. If the amylase content is low, this leads to low dextrin production and poor gas production. This in turn results in inferior quality bread with reduced size and poor crust colour.

However, this is not the only effect of α-amylases. As demonstrated by Pritchard,[54] one of the main effects is the reduction of dough viscosity during starch gelatinization. Gelatinization of non-damaged starch granules starts at 55°C. This leads to amylose leaking out of the granules and initial melting of amylopectin crystallites. These events lead to a sharp increase in dough viscosity, which terminates oven spring. When α-amylases attack gelatinized starch, this will result in a prolonged oven spring and thus larger volume.[48]

Wheat and wheat flour contain endogenous enzymes, of which amylases take an important part. However, the level of α-amylase in some flour is sometimes very low[55] and thus there is

a need for wheat flours to be supplemented with α-amylase. These can be added in the form of malt flour or fungal amylases. Since the 1960s, bakers have supplemented the naturally occurring enzymes in wheat flour to minimize natural differences caused by, for example, weather conditions.

6.2.3 Amylases in bread making

α-Amylases are endoglucanases. This means that they hydrolyze random α-1,4 and α-1,6 linkages. Amylases can act only on damaged or gelatinized starch, since these are susceptible to enzymatic attack. The amount of damaged starch is dependent on wheat variety and especially on milling conditions. Standard UK flour has a higher percentage of damaged starch in order to increase water binding in the dough. Suitable dosages of fungal amylase lead to the desired improvement of dough and the final product. However, extensive degradation of damaged starch due to too high levels of α-amylase leads to sticky dough.

In Figs 6.5(a) and (b) the effect of increasing levels of a fungal α-amylase on volume, crumb structure and stickiness is shown[67] using two different flour qualities.

The volume and crumb structure (manual scores) improve with increasing levels of amylase added to the flour. This effect is seen with different flour types, although the extent of the effect is flour dependent. Even though the positive effects increase with the amylase dose rate, there is an optimum dose level, since the stickiness of the dough also increases, leading, in this case, with flour 1 to an unworkable dough at higher amylase dose levels.

6.2.4 Other amylases

Pullulanase and isoamylase are the two best-known debranching enzymes. Both enzymes are capable of hydrolyzing α-1,6 glucosidic linkages, thereby releasing side chains from the branched amylopectin molecule.

β-Amylase and amyloglucosidase are typical exo-acting enzymes, cleaving α-1,4 glucosidic linkages at the non-reducing end of linear chains in the starch molecule, thereby catalyzing successive removal of β-maltose and β-glucose, respectively.[9,56] β-Amylase is stopped by α-1,6 linkages, whereas amyloglucosidase can bypass the side chains and thus, in theory, can completely degrade starch to β-glucose.[57] These latter four classes of amylases all have a limited effect on dough properties and bread quality.

6.2.5 Anti-staling enzymes

Bread rapidly looses its freshness and is subject to microbial spoilage. Changes in flavour and texture, other than due to microbial spoilage, taking place during storage are commonly called staling. This phenomenon, which makes bread hard and dry, is often attributed to starch retrogradation. Changes that are observed are crumb firming, increased crumb-texture harshness, increase in opacity of the crumb, loss of crust crispness, disappearance of fresh bread flavour and appearance of stale bread flavour.[58] All these factors result in a loss of consumer acceptance of the product.

About 85 million ton of wheat flour is used every year to bake bread. By adding specific agents, such as emulsifiers or enzymes, bread stays fresh longer. It is assumed that 10–15% of bread is thrown away because it not longer fulfils the consumer demands for quality, crumb

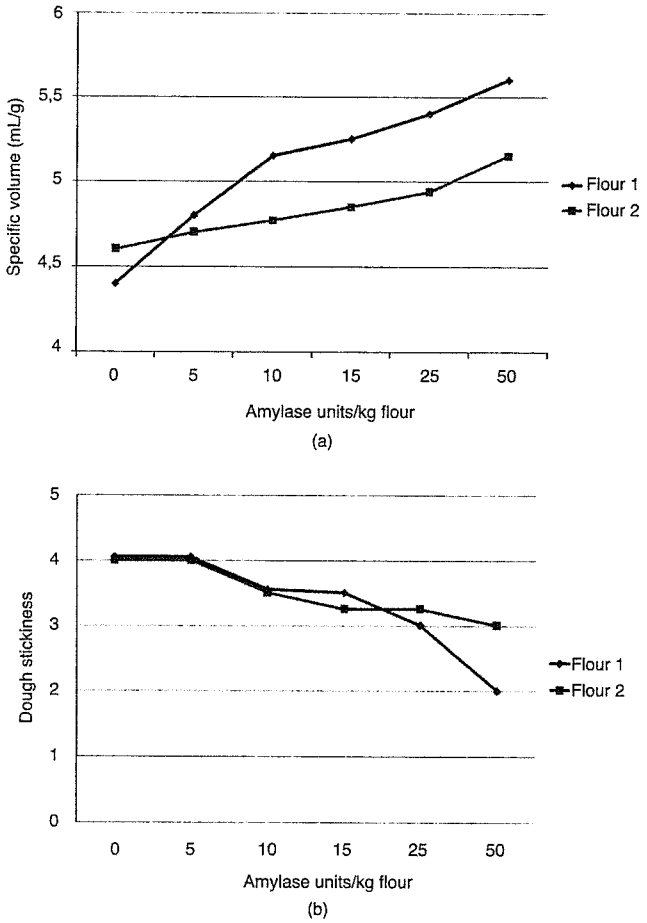


Fig. 6.5 (a) Effects of increasing levels of a fungal α-amylase on specific volume (left) and crumb structure (right). (b) Effects of increasing levels of α-amylase on dough stickiness (score of 5 is not sticky; 0 is most sticky).

softness, taste, etc.; the possibility of keeping it a few days longer fresh could save 2 million ton of flour per year. This is 40% of the annual bread consumption in the US.

Starch retrogradation is seen as the main factor responsible for the observed changes.[32,59,60] However, several authors state that in addition to starch retrogradation gluten, lipids and/or specific dextrins also play important roles in bread staling.[61–66] Furthermore, there are several other factors also having an effect on crumb softness, without necessarily involving starch retrogradation.[67]

- The quality of wheat flour in terms of endogenous enzymes and percentage starch damage has an influence on the total amylase efficiency and thus on the bread quality.

- It is well known that bread volume has a clear relationship with crumb softness. Higher specific volume leads to a softer bread crumb. Fine crumb structure with thin cell walls gives a softer crumb than a coarse structure with thick cell walls.
- Formulation also has a clear influence on staling since any ingredient, like shortening, having an effect on volume will also have an effect on softness.
- Processing has a certain influence on staling as well. Sponge and dough processing gives a different structure and softness compared with straight dough processing. In the same way a twisting step, resulting in a finer and more uniform crumb, will also positively affect softness.
- Finally, storage conditions play a role in staling since storage at lower temperature enhances starch retrogradation and thus will have a clear influence on crumb softness.

In the beginning of this century, the extended shelf life (ESL) concept was introduced in the US. This concept used already existing enzyme technology in order to achieve dramatic improvements in industrial bread making. These improvements were significantly ESLs for industrial bread up till 11 days, strongly reduced stale bread returns and also strongly reduced logistic complexity and costs due to a reduced number of transport routes.[68]

The most visible applications of ESL have been in bread and, to a lesser extent, in snack cakes. However, important improvements are also feasible in other grain-based food products, ranging from cookies via cakes to frozen dough products.

Following these improvements, the demand for even longer shelf lives became obvious. The objectives for further ESL developments can be found within four major areas[69]: texture, flavour, microbial stability and crumb moistness.

With regards to texture, with the use of specific enzymes, like bacterial amylases or intermediate stable maltogenic amylases (see section on amylases) sufficient softness can be obtained. However, sufficient crumb elasticity (or crumb resilience or crumb springiness) is more difficult to achieve. Since it is not fully understood which structures determine resilience, that is gluten, starch, amylopectin, modified amylopectin, etc., it is extremely difficult to find enzymes and/or ingredients which positively influence resilience.

When bread is kept for a longer time, a stale bread flavour develops and the well-known fresh bread flavour and aroma disappear.[70] Removal or masking of this stale flavour is a prerequisite for ESL. Keeping bread soft for a longer period may be feasible by using proper enzymes, but the microbial deterioration has to be suppressed. Increasing the level of propionate in the dough is not an option, due to problems with taste and yeast growth. Longer storage times will result in enhanced evaporation of water from the crumb, via the crust to the outside, thus leaving an unacceptable dry crumb. Specific precautions are needed in order to prevent this.

Further, ESL developments will lead to highly interesting opportunities for industrial baking. Merchandizing, for example showing bread on displays in groceries and supermarkets, baking to inventory – just as is done for cookies and biscuits – and further optimization of distribution and transportation are currently being discussed.

Fungal amylases have limited effect on staling. These enzymes act predominantly on damaged starch, but at the temperature at which starch starts to gelatinize, fungal amylases are already inactivated and thus cannot act on starch when it has become accessible.

Bacterial amylases are much more heat stable and these enzymes have a significant action on gelatinized amorphous starch. Modification of gelatinized starch results in a clear anti-staling effect. However, since bacterial enzymes are extremely heat stable, these

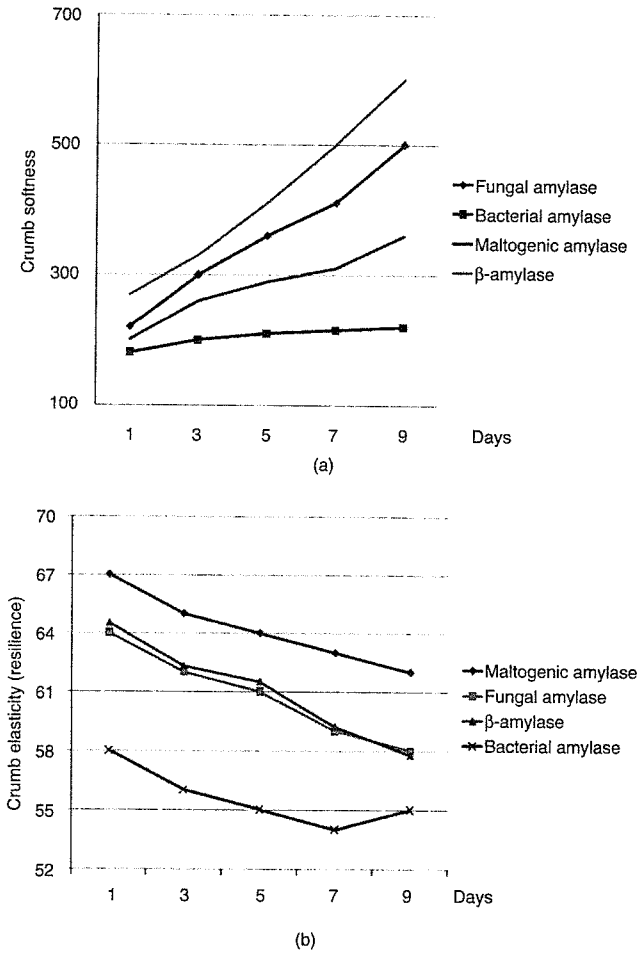


Fig. 6.6 (a) Effect of various starch degrading enzymes on crumb softness as a function of time. (b) Effect of various starch degrading enzymes on crumb elasticity (springiness or resilience).

enzymes exhibit residual enzymatic activity after baking, which can lead to excessive starch degradation, causing a collapse of the bread upon storage after baking. Only at very low dose rates, bacterial amylases can safely be used, but the risk of overdosing remains significant.[71]

Maltogenic amylase (glucans, 1,4-α-maltohydrolase, **EC 3.2.1.33**) produces maltose (and some longer maltodextrines) in the α-configuration. The enzyme is most active between 60°C and 70°C and is able to degrade amylopectin to a greater extent than fungal amylases or β-amylase.

In the above graph (Fig. 6.6(a)), the effects of several different amylases on crumb freshness are shown.[72]

Table 6.3 The effects of various amylases on staling.

Enzyme	Mechanism	Thermostability	Softening	Springiness
α-Amylase (A. oryzae)	Mainly endo	Low	+	Very limited
α-Amylase (A. niger)	Mainly endo	Intermediate	+	Little
α-Amylase (B. amyloliquefaciens)	Endo	High	++++	Negative
α-Amylase (Maltogenic)	Exo and endo	Intermediate	++++	Positive
β-Amylase (e.g. from wheat)	Exo	Low	+	Little

As can be seen in Fig. 6.6, maltogenic intermediate stable amylase has a significant effect on crumb softness as function of storage time. A specific bacterial amylase even has a greater effect on softness, but this same enzyme completely ruins the crumb elasticity (Fig. 6.6(b)), whereas the maltogenic enzyme gives the crumb a relatively good springiness, even after prolonged storage.

Fungal amylase and β-amylase have a very limited effect on crumb softness and on crumb resilience. The effects of these enzymes are comparable with effect of distilled monoglyceride.

The maltogenic amylase has a thermostability which is in between those of fungal amylase and bacterial amylase.[73] Therefore the enzyme is able to reduce retrogradation of amylopectin. It can hydrolyze glucosidic linkages in gelatinized starch before it is inactivated during the baking process. Since the enzyme is inactivated at the end of the baking process, it does not excessively hydrolyze starch. Apart from the optimal thermostability of this enzyme, it has other benefits compared with fungal or bacterial enzymes.[74]

The enzyme can degrade amylase and amylopectin into maltose and longer maltodextrines, and in doing so it does not need an unblocked non-reducing end. This indicates that the enzyme is also capable of attacking starch through an endo-type mechanism. Besides that, these maltodextrines are also assumed to have an anti-staling effect by specifically blocking interactions between starch and gluten.[62,63] The effects of various amylases on staling are summarized in Table 6.3.

6.3 XYLANASES

Numerous studies have been performed to demonstrate the positive effects of pentosans-modifying enzymes, which are presented by industry as pentosanases, xylanases, arabinoxylanases and/or hemicellulases, here further referred to as xylanases.[29,75–78] The way these enzymes are considered to have their effect is by reducing the water binding of WU-AX and by solubilizing WU-AX and later also WE-AX into smaller molecules.[79,80] Another effect ascribed to xylanases is to offset reduced gluten coagulation caused by pentosans by hydrolyzing the pentosans to an extent whereby this effect is not longer occurring.[81] Hamer[79] reported that the use of xylanases in a batter significantly improved gluten coagulation. The resulting gluten also was shown to have a much better bread-making quality. This effect was explained by the absence of any detectable pentosans in the remaining gluten, whereas normally 2–3% pentosans were attached to gluten. These gluten-linked pentosans were considered to have a steric hindrance effect on gluten coagulation.[82] Currently, industrial xylanases are sold to the starch industry for processes whereby gluten and starch are separated, which are based on this principle.

6.3.1 Classification

Glycoside hydrolases have been classified into 93 families[52] based on sequence homologies, which reflect structural and mechanistic features.

Xylanases (endo-1,4-β-D-xylanase, **EC 3.2.1.8**) can be classified in at least three ways. The first is based on molecular weight and pI.[83] They either have HMW or LMW and have either a high (basic) or low (acidic) pI. The second is based on crystal structure. This can be derived indirectly by a determination of DNA sequence. Endo-β-1,4-xylanases are generally classified in families 5, 8, 10 (formerly family F), 11 (formerly family G), 16, 26 and 43.[52] However, the majority belong to families 10 and 11 and are found to frequently have an inverse relationship between their pI and molecular weight. The family 10 xylanases are generally larger and more complex than the family 11 xylanases. The third classification is based on kinetic properties, substrate specificity or product profiles. Virtually all xylanases are 'endo' acting, as readily determined by chromatography, but the more detailed determination of kinetic properties, measuring the relative reaction rates on various substrates and determining the kinetics of intermediate product formation, is much less common.[84]

Unfortunately, only very few studies have been performed in order to relate sequence or structural family classification to action patterns, substrate specificity or functionality (e.g. in bread making).

Family 10 xylanases occasionally exhibit endocellulase activity; they generally have a higher molecular weight, and they occasionally will possess a cellulose-binding domain. Also in general these enzymes are considered to be less specific.[85]

Members of family 10 (including all plant xylanases, such as xylanase from cereals) will act on both PNP-xylobiose and PNP-cellobiose (artificial chromogenic substrates); however, the overall catalytic efficiency on PNP-xylobiose is about 50 times higher. This suggests that family 10 enzymes act mainly on xylan. Family 10 xylanases are capable of attacking the glycosidic linkages immediately next to branching points and the endoxylanases require two unsubstituted xylopyranosyl residues between the branches.[85]

Even though all xylanases are endo acting, they show variations in their product profiles. Some enzymes generate predominantly xylose and xylobiose and others predominantly (or exclusively) form xylotriose or a range of other/higher oligosaccharide products. Family 10 xylanases have relatively HMWs, and they tend to form oligosaccharides with a low degree of polymerization (DP). These xylo-oligosaccharides can be further degraded by β-xylosidases (**EC 3.2.1.37**), which remove xylose from the non-reducing end of the polymeric xylan residue.

Family 11 xylanases are true xylanases. They do not have cellulase activity; they consistently exhibit a LMW, and they can have either a high or low pI. They are formed by both bacteria and fungi. The positions of many amino acids are essentially identical in the family 11 xylanases from bacterial (*Bacillus circulans*) and fungal (*Trichoderma harzianum*) origins. Thus, there has been a tremendous conservation of the basic structure of the catalytic site of family 11 xylanases during evolution.[86,87] This is remarkable when considering the differences in functionality between the two classes of enzyme in bread making. Family 11 endoxylanases require three consecutive unsubstituted xylopyranosyl residues and thus also in this way can be distinguished from family 10 xylanases.[88]

Some xylanases belonging to GH families 5, 8 and 43 have been identified[89–91] (see Ref. [52] with the EC code for xylanase 3.2.1.8. for an overview). These have not been studied in any detail and their bread-making potential has not been elucidated in any detail. Among these, one example is the xylanase from Pseudoalteromonas haloplanktis TAH3a belonging to glycoside hydrolase family 8.[92–94] This enzyme is a typical psychrophilic enzyme and presents a high catalytic activity at low temperatures. It is not homologous to family 10 or 11 xylanases, but has 20–30% identity with glycoside hydrolase family 8 members (formerly family D), a family that comprises mainly endoglucanases, but also lichenases and chitosanases.

6.3.2 Mechanism

The mechanism of action of xylanases in bread preparation is still not clearly elucidated. Many types of hemicellulase preparations have been used for the applications mentioned above, and are commercially available. They are produced by microbial fermentation using various microorganisms as enzyme sources. Many of these enzymes are produced by genetically modified microorganisms. All documented commercial uses of xylanases relate to enzymes belonging to either glycoside hydrolase family 10 or family 11, as defined previously. Examples of commercial xylanases are the xylanases from *Bacillus* sp., *Trichoderma* sp., *Humicola* sp. and *Aspergillus* sp.

It has been assumed for a long time that, depending on the application, preferential attack of WE-AX or WU-AX is needed, whereas activity towards the other fraction is not desired.[29] In bread making, endoxylanases that have a preference towards WU-AX have been considered beneficial.[29,95] The difference in substrate selectivity is therefore an important parameter in developing and selecting proper xylanases.[26] Comparison of selectivities and activities of a Bacillus xylanase and an Aspergillus xylanase showed that the Bacillus enzyme has a clear preference for WU-AX, whereas the Aspergillus xylanase more readily hydrolyzed WE-AX.[96] Nevertheless, both enzymes have a certain (although not the same) positive effect on bread making, confirming the findings of Wang[24] that both WE-AX and WU-AX affect gluten network formation in a similar negative manner. This means that hydrolyzing either one of these components can have a positive effect.

Figure 6.7 shows that the positive effect of xylanases is more pronounced when bread is made with low protein flour.[97]

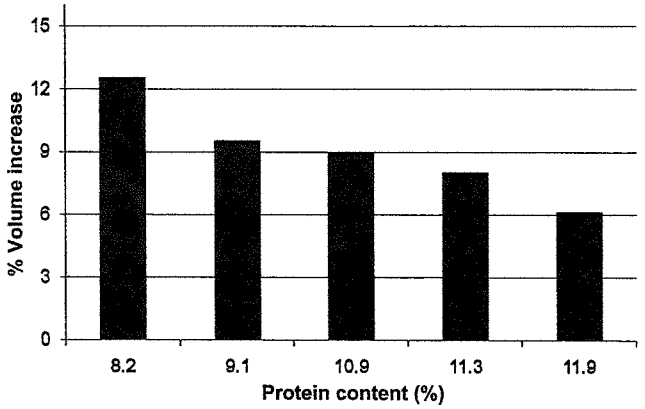


Fig. 6.7 Effect of a xylanase, in terms of per cent volume increase, on flour types with varying protein contents.

Again this is in line with the results from Wang.[24] A negative influence on gluten formation will be more difficult to handle in low protein flour than in high protein flour. Offsetting the negative effect will therefore be more effective in low protein flour.

6.3.3 Xylanases in bread making

Xylanases are broadly used in bread making, and depending on the application there is generally an appropriate xylanase or a mix of different xylanases that gives the desired effects in terms of dough-handling properties;[29] stability and oven spring[98] and volume.[95] This immediately indicates that there is not one single xylanase giving all desired effects in any application, but that the xylanase type(s), usage and dose rates need to be optimized in each case.

In spite of all research that has been done on xylanases and in spite of the wide acceptability of this type of enzymes, there is still no complete understanding of the mechanism and effects of different xylanases. This becomes obvious from a range of experiments (results not shown) in which four different xylanases (a monocomponent genetically modified organism (GMO)) – Aspergillus xylanases, an Aspergillus xylanase produced by solid state fermentation (thus having a wide range of side activities), a bacterial (Bacillus) xylanase and a fungal (Trichoderma) xylanase – are compared for different applications. Each of these four enzymes, in spite of similar dose rates, has a different effect on dough and bread properties when tested in different applications, varying in mixing time, water addition, mixing type and bread type. This indicates that there is not one single xylanase which performs equally well under all circumstances. Bakeries and bread improver companies need to establish optimal dose rates and optimal xylanase blends for each application and they can do this by trial and error only. There is no way to predict the performance of a xylanase.

6.4 LIPASES

Since a few years lipases (glycerol ester hydrolases, **EC 3.1.1.3**) and phospholipases (A2 and A1 type; **EC 3.1.1.4** and **EC 3.1.1.32**, respectively) are recognized as an additional tool for improving bread-making properties and in particular for their strong, positive effects on dough conditioning and dough characteristics.

Lipases hydrolyze ester bonds of acylglycerols, yielding mono- and diacylglycerols, free fatty acids and, in some cases, also glycerol. Lipases preferably hydrolyze ester bonds at the sn-1 and sn-3 position of the glycerol molecule.[99] Lipases usually function at lipid-air or lipid-water interfaces and their activity is sharply increased by the presence of organized lipid structures, which are normally found at such interfaces.[100,101]

Lipases can generally be divided into four groups, according to their specificity: substrate specific lipases, regioselective lipases, fatty acid specific lipases and stereospecific lipases.

6.4.1 Mechanism

The structure of a bread dough can be seen as a foam structure. The individual gas cells are separated by a continuous gluten film in which the starch granules are also embedded.

Bread-making quality is largely determined by the gas cell stability.[102] The distribution of the gas bubbles in the dough and also their size are largely determined by the flour quality, the ingredients and the mixing conditions. Some studies have focused on the role

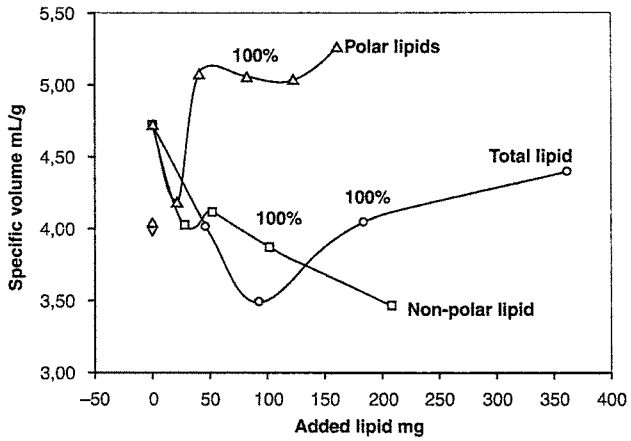


Fig. 6.8 Effect of addition of various wheat lipids on bread volume (adapted from Ref. [97]). Guy, R.C.E.; Sahi, S.S. (2002) Comparison of effects of xylanases with fungal amylases in five flour types. In: Recent advances in enzymes in grain processing. Courtin, Veraverbeke and Delcour eds. Lab. Food Chem., Catholic University Leuven, Belgium.

of gluten proteins on the gas holding capacity,[103] but there is also evidence that a lipid film surrounding the gas cell is also contributing to the gas cell stability.[47,104] Surface-active materials (emulsifiers) are able to counteract instability of gas cells. In fact these ingredients prevent coalescence and disproportionalization of gas cells by stabilizing the interface.

Initially, positive effects of lipase action were explained by assuming that lipases break down non-polar lipids, such as triglycerides, thereby removing a negative component from the dough. As can be seen in Fig. 6.8,[105] non-polar lipids generally have a negative effect on bread-making properties. It is well-known fact that triglycerides do not form stable monolayers at the lipid-air interface, and following this logic it is understandable that breakdown of triglycerides was considered to be positive. The 1,3 specific lipases were also assumed not to attack the polar lipids, which seemingly have a positive effect on bread-making quality.[106]

The current generation of lipases hydrolyze both polar lipids (see Fig. 6.9) as well as non-polar lipids. The resulting products show clear structural resemblance to well-known emulsifiers, such as DATEM and SSL. In that way, lipase action results in the stabilization of gas cells. However, there is a clear difference in effects of specific 1,3 specific lipases and lipases having other specificity, that is towards polar lipids.

Figure 6.9 shows the effect of lipase and phospholipase action on the structure of specific lipids. DGDG (digalactosyl-diglyceride) is converted into DGMG (the monoglyceride variant), and phosphatidylcholine (lecithin) is converted into lysophosphatidylcholine.

6.4.2 Lipases in bread making

1,3 Specific lipases are claimed to improve dough-handling properties, to increase dough strength and stability, to improve dough machinability and to increase oven spring. Besides this, such lipases also improve crumb structure and crumb whiteness.[107] The first generation of lipases in baking were almost exclusively of this type and were claimed to be alternatives to chemical dough strengtheners and emulsifiers. However, the technical and commercial

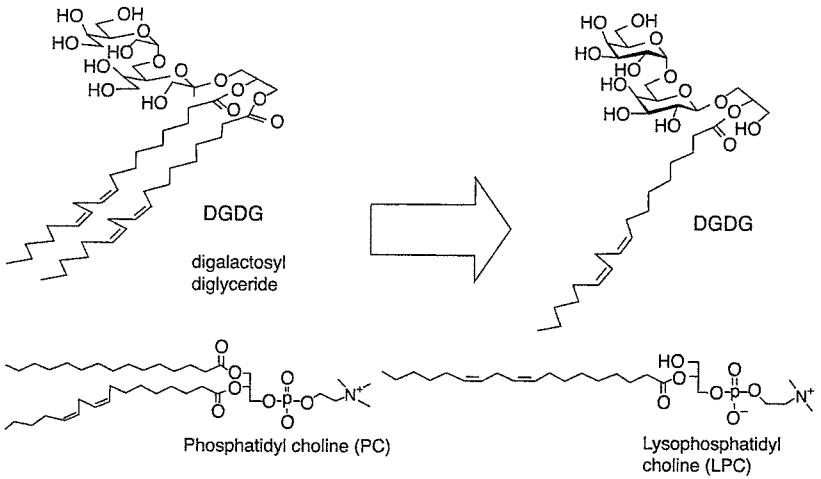


Fig. 6.9 Molecular structure of various wheat lipids before and after lipase hydrolysis (adapted from Ref. [105]).

benefits were limited. The second generation were enzymes with much broader specificity, acting also on polar lipids. These enzymes exhibited phospholipase activity as well as lipase activity. Phospholipases act much more like emulsifier replacers.[108] Both types of lipases were shown to give an increase in surface pressure of gas cells, although the effect of phospholipases was much more pronounced.[108] A similar increase in surface pressure is also obtained by addition of DATEM. Increased surface pressure leads to a better distribution of more stable, smaller gas cells. This results in finer, more silky crumb structure with optically whiter colour, better dough-handling properties and, to a certain extent, a larger loaf volume.[109] However, the surface pressure alone cannot account for the positive effects of the enzymes or of the emulsifier.[108] Further research is needed in order to elucidate the reaction mechanism of the various lipases, their breakdown products and the effects of these products in the bread-making process.

A third generation lipase is currently entering the baking market. These enzymes are protein engineered in order to give a better effect in high speed mixing and no-time dough processes. Especially in these processes, the first and second generation lipases were not very successful. Furthermore, the third generation lipases that become available now have a lower affinity for short chain fatty acids, leading to lower release of such fatty acids, which result in lower risk for off-flavour formation upon prolonged storage of the baked goods and upon use of butter or milk fat in baked products.

Lipases are also claimed to have a direct effect on the gluten network.[110,111] This can partly be explained by the generation of free fatty acids, which can be oxidized by endogenous lipoxygenase, leading to an increased oxidation potential which in turn may positively affect gluten network formation. Furthermore, lipase may influence the interaction between gluten proteins and lipids and the interactions between starch and lipids. Especially, the anti-staling and crumb softening effects of lipases are said to be the results of amylose–lipid complex formation.[112,113] These effects could not be obtained by simply adding monoglycerides to

the dough, indicating that the amount of emulsifier-like structures formed due to the lipase action is not fully explaining positive lipase effects.[108]

Concluding, it will be clear that there is still a lot of uncertainty about the mechanism of lipase action in dough and in bread making. More research with even more specific lipases is needed in order to reach well-defined conclusions.

6.5 OXIDOREDUCTASES

Oxidoreductases are widely distributed among microbial, plant and animal organisms. These enzymes catalyze the exchange of electrons or redox equivalents between donor and acceptor molecules. This occurs in reactions involving electron transfer, proton abstraction, hydrogen extraction, hydride transfer, oxygen insertion or other key steps.[114] In general two half reactions, one oxidative and one reductive, take place and at least two substrates (one reducing and one oxidizing) are activated or transformed.

To accomplish this physiological function, oxidoreductases have various redox-active centres.[115] Common redox centres include amino acids, metal ions, metal complexes (e.g. Fe-S clusters; heme cluster) or coenzymes (e.g. FAD, NAD, pterin, PQQ).

Many oxidoreductase substrates, such as carbohydrates, unsaturated fatty acids, phenolics and thiol-containing proteins, are important components of wheat flour. Their modification by oxidoreductases may lead to new functionalities, quality improvements and/or cost reduction.

6.5.1 Classification

Oxidoreductases can be classified according to their amino-acid sequence, their three-dimensional structure or their application, that is type of catalysis and/or cofactor dependence.[114] In the latter classification according to application, four groups can be distinguished:

- Oxidases
- Peroxidases
- Oxygenases
- Dehydrogenases/reductases

Within each of these groups, various subtypes exist, mainly dependent on differences in active centres. Within the baking industry, several oxidases have been investigated and some have been commercialized.

6.5.2 Oxidases in baking

In bread making, bread improvers and dough conditioners are widely used and accepted. The main action of these agents is to help rebuild the gluten network and the GMP, in order to improve texture, volume, freshness and also dough machinability and stability. Dough conditioners are specifically meant for gluten strengthening. Gluten strengthening results in improved rheological and handling properties of the dough. Such conditioners also have a long history in bread making and are well known. Non-specific agents, such as iodates, peroxides, ascorbic acids, potassium bromate and azodicarbonamide, all have a

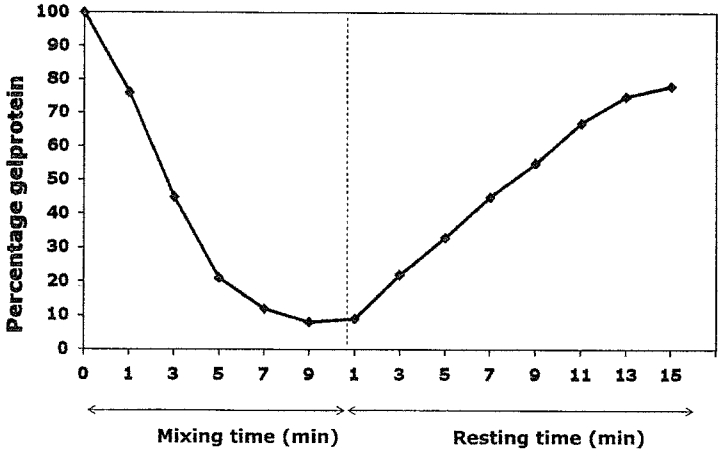


Fig. 6.10 Breakdown and rebuilding of gel protein during dough mixing and dough rest.[16] Weegels, P.L.; Hamer, R.J.; Schofield, J.D. (1997), Depolymerization and repolymerization of wheat glutenin during dough processing II. Changes in composition. J. Cereal Science. 25: 155–263.

gluten strengthening effect by inducing the formation of protein–protein bonds that strengthen the protein network and thereby strengthen and stabilize the dough.[116]

The glutenin sub fraction, which is unextractable in SDS solution (gel protein or GMP) and which is highly correlated to various quality parameters of bread, is changing during bread making. During dough mixing, GMP partly depolymerizes, which leads to more SDS-soluble glutenin. During resting, these extractable proteins repolymerize, thus increasing GMP again (see Fig. 6.10).[16] This process of disaggregation and rebuilding can be influenced by longer or shorter mixing; but the whole process is catalyzed by oxidoreduction reactions.

It is generally accepted that the properties of dough and its three-dimensional protein network are dependent on the arrangement and number of disulphide bonds and sulfhydryl groups of the protein. The vital contribution of disulphide bonds to dough stability has been shown by rheological studies,[117] and it is specifically this group of reactive groups that are targeted by oxidative agents or oxidases.

After mechanical development of the gluten network, the three-dimensional protein structure needs to be stabilized by oxidants. Small amounts of oxidizing reagents, such as potassium bromate or dehydroascorbic acid, improve the dough handling and baking characteristics of wheat flour; loaf volume increases and bread crumb improves as well.[118] Bromate is assumed to oxidize LMW, SH-containing peptides (glutathione) into disulphide bonds.[119] On the other hand, a small amount of cysteine or reduced glutathione sharply increases the extensibility of dough. Both the viscous and elastic component of dough deformation are increased by addition of these reducing agents.[117]

In general, due to the great number and also the complexity of oxidoreduction reactions occurring during bread making, the effects of oxidative reagents are only poorly understood.[120–122]

Oxidizing effects can also be obtained by using enzymes instead of chemical dough conditioners. In Table 6.4, a number of oxidases are listed which are currently used in commercial bread making or which have been investigated specifically for this purpose. As

Table 6.4 Oxidases investigated and/or commercialized for bread-making applications.

Enzyme	EC number
Glucose oxidase	1.1.3.4
Hexose oxidase	1.1.3.5
Pyranose oxidase	1.1.3.10
Sulfhydryl oxidase	1.8.3.2
Glutathion oxidase	1.8.3.3
Glutathione dehydrogenase (DHA reductase)	1.8.5.1
Diphenyl oxidase (catechol oxidase)	1.10.3.1
Laccase	1.10.3.2
Ascorbic acid oxidase	1.10.3.3
Peroxidase	1.11.1.7
Glutathion peroxidase	1.11.1.19
Lipoxygenase	1.13.11.12
Tyrosinase (polyphenol oxidase)	1.14.18.1

can be seen, there is quite some variety in enzymes from different subclasses, which all are claimed to give a beneficial effect in bread making. The reactions of most of these enzymes in wheat dough are schematically shown in Fig. 6.11.

Glucose oxidase and hexose oxidase most likely catalyze their reactions following a similar reaction mechanism. Glucose, preferably in the β form, is oxidized by glucose oxidase to form gluconolacton, which is immediately converted into glucuronic acid. For

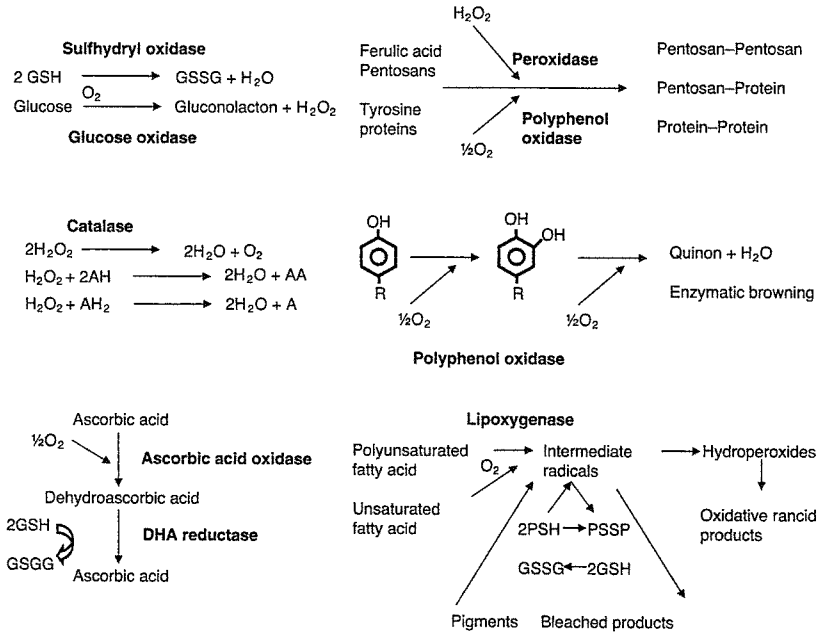


Fig. 6.11 Oxidation reduction systems in wheat dough.

the reaction the presence of oxygen is required. Oxygen serves as an electron acceptor. In this reaction hydrogen peroxide is formed. One explanation of the reaction mechanism is that hydrogen peroxide, in the presence of endogenous peroxidase, naturally occurring in flour, promotes the oxidation of sulfhydryl (—SH) groups to disulphide (S—S) bridges in the gluten network, as well as the formation of a gel from the water-soluble pentosans.[123,124] Another explanation is that hydrogen peroxide reduces the level of reduced glutathione, which normally has a weakening effect on the gluten network formation.[125]

The increase in gluten network results in increased strength of the gluten structure in the dough. This leads to improved dough stability, reduced dough stickiness and improved dough machinability. These improvements in turn lead to increased volume, improved crumb structure and better softness of the baked product. However, there are also studies which do not support these theories, since no evidence could be found for increased or changed gluten structures.[126] The main difference between glucose oxidase and hexose oxidase is that the latter can use various monosaccharides and even oligosaccharides as a substrate.

Lipoxygenase converts polyunsaturated fatty acids, more specifically those containing a *cis*, *cis*-1–4-pentadiene moiety, to fatty acid peroxy radicals. This reaction also requires the presence of oxygen. The free radicals react further to yield monohydroperoxides with conjugated double bonds and these compounds can react further with a wide variety of products.[121,127]

In bread making, lipoxygenase from soya bean flour has been used for decades not only for its bleaching effect, resulting in a whiter crumb,[128] but also for its improving effect on dough rheology (viscoelasticity), on mixing tolerance, loaf volume and on the gluten stability.[129–132] The hydroperoxides react with the naturally occurring yellow carotenoid pigment in wheat flour, leading to a reduction of the yellow colour. Furthermore, it has been claimed that lipoxygenase has a direct oxidizing effect on gluten formation.[121,133,134] This effect cannot be ascribed to hydroperoxides, since addition of lipid hydroperoxides did not show any effect.[135,136] Most of the oxygen uptake by wheat dough during mixing is due to the oxidation of free and esterified polyunsaturated fatty acids (PUFAs).[137–141] The improving effect of lipoxygenase may thus be due to the oxidation of gluten proteins through the co-oxidation of accessible thiol groups of the gluten protein by the enzymatically oxidized lipids.[142–146]

Wheat flour itself also contains lipoxygenase, but the activity of this enzyme is confined to free linoleic acid, linolenic acid and monoglycerides containing these fatty acids.[147]

Currently there are no other commercial sources of lipoxygenases other than enzyme-active soya bean flour and, to a lesser extent, flour from other beans (e.g. faba beans). With the current trend towards liquid bread improvers, either water based or oil based, the usage of soya bean flour is gradually reducing as a consequence of the limited solubility of soya bean flour. This increases the necessity for a microbial lipoxygenase. However, it seems extremely difficult to find a lipoxygenase with exactly the right specificity. Also in soya beans there are three distinct isoforms of lipoxygenase (LOX1, LOX2 and LOX3), of which only LOX1 and LOX3 have a positive effect on loaf volume. LOX2 is mainly responsible for undesirable aroma formation in bread dough.[148]

Polyphenol oxidases (PPO) are normally involved in enzymatic browning reactions. Enzymatic browning is the enzymatic oxidation of phenols leading to the formation of pigments. The colour of those pigments varies widely in colour and intensity.[149]

Several names are in use for PPO, including phenolase, creolase, tyrosinase, diphenolase, catecholase, laccase, etc. According to the official nomenclature, two kinds of enzymes are classified under the trivial PPO name. The first group, catechol oxidases (or diphenol oxidase) (**EC 1.10.3.1**) catalyzes two distinct reactions in the presence of oxygen: the hydrolxylation of monophenols into *o*-diphenols and the oxidation of the *o*-diphenols into *o*-quinones. Both reactions require oxygen. The second group, laccases (**EC 1.10.3.2**) oxidizes both *o*-diphenols and *p*-diphenols, thereby forming the corresponding quinines. However, laccases can also act on monophenols. A third group of enzymes exists, which is contributing to the confusion about nomenclature of those enzymes: tyrosinase (or polyphenol oxidase, **EC 1.14.18.1**). The latter also catalyzes phenol oxidation, but by a different mechanism, involving two electron transfers.[150] Oxidation of tyrosine moieties in proteins can also lead to the formation of new covalent bonds, for example with cysteine groups, which in turn leads to changed rheological properties.[151–153] In spite of positive findings,[154,155] there are currently no commercial tyrosinases available for baking.

Laccases (**EC 1.10.3.2**) are copper containing enzymes that have been extensively tested for baking applications but this has not resulted in a successful commercial product. Laccases are claimed to increase dough stability and dough strength and to reduce dough stickiness.[65] Both laccases and peroxidases (cf. below) catalyze the oxidative gelation of arabinoxylans in model systems.[36] In bread systems, it has been demonstrated that laccase reduces the extractability of arabinoxylans due to cross-linking of AX chains as a result of dimerization of FA residues.[23] It has also been suggested that these enzymes only catalyze the formation of a sugar network and not the formation of a gluten network.[156] It is more likely, though, that both cysteine and tyrosine residues are also involved in oxidative cross-linking reactions.[33] In this reaction, laccase is assumed to catalyze the formation of thiol radicals as a result of the formation of phenoxyl radicals. Finally, proteins may also be coupled to arabinoxylan chains by FA moieties and tyrosine or cysteine residues.[27,157]

Currently, several industrial laccases are for sale (for use in textile, in juice, in brewing), but none of them has been commercialized directly for the baking industry and thus there are little or no laccase sales for baking applications.

Peroxidases (**EC 1.11.1.7**) are also able to catalyze the oxidation of phenolic groups. Peroxidases use hydrogen peroxide as electron acceptors and can use a wide variety of substrates, leading to the formation of radicals which can react further, non-enzymatically, with other substrates.[158] In this way, peroxidase shows a dough strengthening effect, leading to improved volume and crumb characteristics.[158] Peroxidase causes oxidative gelation of soluble flour arabinoxylans (pentosans). This oxidative gelation of pentosans by hydrogen peroxide is ascribed to FA residues. One possible mechanism is through dimerization of FA residues on adjacent arabinoxylan chains.[159] Proteins also participate in this gelation, since the gel fraction contains around 25% protein. In this case, the mechanism is by coupling FA residues to tyrosine or cysteine residues in protein chains.

Peroxidases are commercially available, although the only food grade version is derived from plant material (soya bean hulls), whereas microbial-derived peroxidases are commercialized for non-food applications, such as the production of textile materials.

Sulfhydryl oxidase (**EC 1.8.3.2**) catalyzes the formation of disulphide bonds from a variety of thiol groups. This enzyme was originally used for removal of cooked flavours in UHT milk.[160] It was speculated that sulfhydryl oxidase would give a similar effect as that of chemical oxidants[161] and that chemical oxidants would therefore lead to formation of disulphide bonds between protein chains. For this reason the enzyme has been extensively studied for its effect in bread making. However, it was found that the enzyme had no effects on loaf volume, dough strength, dough stability and mixing tolerance.[161] The conclusions were that sulfhydryl oxidase had no or very limited affinity for thiol groups in protein chains, but only for small thiol containing molecules.

Nevertheless, the enzyme has been commercialized since it was claimed that there was a positive synergism with glucose oxidase[162] and it has been available for a number of years. Nowadays it has been withdrawn from the market for baking enzymes.

Amino acid oxidase (EC 1.10.3.3) and Dehydroascorbic acid dehydrogenase (EC 1.8.5.1) are both mentioned in relation to the oxidation and reduction of ascorbic acid (AA) in wheat dough and the concomitant oxidation of glutathion. The improving effect of AA is mediated through its oxidation in dehydroascorbic acid (DHA). The latter compound is able to oxidize two thiol groups into a disulphide bridge.[163] Although AA oxidase has been characterized in wheat flour by many researchers,[164–166] the possible oxidation of AA by other enzymes[167] or by non-enzymatic reactions[165, 168] cannot be excluded.

Reduction of DHA into AA with the concomitant formation of disulphide bridges from thiol groups is an enzymatic reaction, especially when the thiol groups are from glutathion.[134, 169] Glutathion DHA reductases are present in wheat flour and wheat bran.[170] This enzyme is specific for glutathion and inactive against cysteine and cysteine-containing peptides.

In spite of the clear effects these enzymes have on the oxidation reduction systems in wheat dough, none of them has been commercialized.

Concluding remarks on oxidases may be that in spite of extensive research done globally on oxidases, the commercial success is so far rather limited. One of the reasons that has frequently been given is the requirement of oxidases for molecular oxygen. Since the amount of oxygen in a dough is limited and also taken up by yeast, this could explain the limited success of oxidases. However, incidentally, very positive results have been obtained with oxidases.[171] This, in combination with the lack of complete mechanistic understanding of the oxidation processes occurring in dough, leads to the conclusion that positive oxidase effects must be possible. Most likely the right oxidases still need to be discovered and commercialized.

6.6 PROTEASES

Proteolytic enzymes, also referred to as proteases, proteinases and peptidases, catalyze the hydrolysis of peptide bonds in proteins. A wide variety of proteases exist in nature; in plant material, animal tissue and in many different microorganisms. Commercially, proteases are by far the largest group of enzymes sold for a wide variety of applications.

Commercial proteases can be of cereal (or other plant materials), animal, fungal or bacterial origin. In contrast to, for example, amylases, proteases do not differ much in terms of heat stability. They differ widely, however, in terms of pH dependence and even more in terms of catalytic specificity.

6.6.1 Classification

Proteases can be divided into two subclasses: endoproteases and exoproteases. The first group hydrolyzes peptide bonds of proteins in the interior of the polypeptide chain, thereby generating smaller peptides and sometimes even free amino acids. Endoproteases can be further subdivided into four groups:

Serine proteases (**EC 3.4.21**...)
Cysteine proteases (**EC 3.4.22**...)
Aspartic proteases (**EC 3.4.23**...)

Metalloproteases (**EC 3.4.24**...)
(... indicates that various sub-types exist.)

This classification is based on the catalytic mechanism of the enzymes and specific functional groups or molecules involved in the catalytic action.

Exoproteases or peptidases cleave the protein chain at the end, thereby generating free amino acids and sometimes even small peptides. Peptidases are usually subdivided into two classes, each with various subclasses:
Carboxypeptidases (**EC 3.4.16**..., **EC 3.4.17**..., **EC 3.4.18**...)
Aminopeptidases (**EC 3.4.11**...)

This subdivision is based on the specificity of the enzymes, that is from which side of an amino acid the peptide bond is hydrolyzed.

6.6.2 Proteases in baking

The most important functional component in wheat flour is gluten. Therefore, anything that influences or modifies the gluten network or the individual gluten proteins, and thus the ability to form a network, will have a strong influence on the dough and final bread quality. Degradation of gluten proteins has an immediate effect on the covalent interactions in the gluten network.

Proteases have a long history in bread making and were traditionally used to treat 'bucky' dough resulting from overly strong and too elastic flours.[172] Originally the aim of protease addition was to improve softness, dough-handling properties and dough machinability.[172, 173] However, proteases have more functional effects. Functional effects of proteolytic enzymes are reduction of mixing time; improvement of dough machinability; improvement of gas retention due to better extensibility; improved pan flow in bun and roll production; improvement of grain and crumb texture; improved water absorption; improved colour; and improved flavour.[174, 175]

When achievement of such changes in dough characteristics is the purpose of protease addition, it makes sense to add an endo-acting protease. Hydrolyzing internally located peptide bonds will have a much more dramatic rheological effect than a removal of a terminal amino acid by the action of an exo-acting peptidase.[176]

Apart from effects coming from added enzymes, there are numerous reports on effects from proteases coming from infections on wheat,[177, 178] from lactobacilli used in sourdough preparations[179] and from endogenous enzymes, for example in sprouted wheat.[172, 180] Infections can result in significant degradation of endosperm protein, lower amounts of storage protein, lower dough consistency, increasing resistance to extension and lower loaf volume.

6.6.2.1 *Bread flavour*

Limited proteolytic hydrolysis, as seen in sourdough, has a positive effect on bread flavour. Bread crumb flavour is to a large extent formed by 2-acetylpyroline, whereas bread crust is to a large extent determined by a limited number of volatile compounds formed during fermentation.[181] Many of these volatiles originate from fatty acid oxidation or from microbial metabolized amino acids. In sourdough, the level of free amino acids is much less than in normal bread[182] and the microbial conversion of certain amino acids (ornithine, leucine, phenylalanine) leads to improved bread flavour.[77, 181] Enhanced proteolysis in sourdough leads to the formation of higher levels of free amino acids and thus to improved sensorical

properties compared with yeasted dough.[183] During this proteolysis, mainly HMW glutenin subunits are broken down, leading to increased glutenin solubility and reduced ability for network formation.[184]

The effects of proteases are strongly dependent on bread-making methods used, on flour quality and on the presence of other functional ingredients. In a short process, the effects of a protease on mixing requirements were found to be negligible, whereas effects on volume and bread score were strongly dose dependent: at low dose rate there was a clear improvement in volume and bread score, whereas, at especially higher dose rates, overall bread score greatly reduced.[185] On the other hand, in a sponge and dough system, proteases greatly reduced mixing time.[186] In shelf stable bread systems, proteases reduced the firmness of the crumb more than other enzymes and also reduced moisture migration.[187] In both cases, effects of these proteases on dough volume were limited, but proteases had a strong reducing effect on crumb firmness and bread score.[185,186]

Most of the effects mentioned are derived from modification of the gluten proteins. Limited proteolytic hydrolysis with a degree of hydrolysis (DH) of 0–5%, leads to an increase in gluten solubility, in combination with an improvement of the emulsifying and foaming properties of the gluten. The released soluble peptides had decreased functional properties.[188]

Proteases can have major disadvantages. The action of the proteases is not limited in time, they continue after mixing and weaken the dough structure in time. This phenomenon increases the risk of weakening the dough and increases the stickiness of the dough. Sometimes their action is even enhanced by the pH drop during fermentation. The use of proteases in baking requires strict control of the bulk fermentation and proofing conditions of the dough. Almost all proteases are inactivated during baking. Especially care should be exercised when using neutral Bacillus proteases and papain, which should be dosed very carefully as overdosing will slacken the dough too much. This may result in dough collapse before baking or a lower bread volume and a more open crumb structure. Particularly in Europe, where the flours are weaker than in the US or Canada, the risk of overdosing a protease is very much present. Furthermore, proteases also increase stickiness because by hydrolytic action, water is released from the gluten. This means that in practice proteases are little used in bread making in Europe.

6.6.2.2 *Freshkeeping*

Proteases have also been investigated in relation to freshkeeping. As mentioned above, there are several ingredients (emulsifiers, fats, mono- and oligosaccharides) and processing tools (specific amylases) known to retard bread staling. As there was not always a good correlation between starch structure and staling, other flour constituents were also investigated. The role of flour proteins in the crumb firming process has been studied but it was found that they were less important than starch.[189]

Nevertheless, upon using an intermediate thermostable or thermostable protease in baked goods a pronounced effect on crumb softness and on retarding the staling of baked products was found.[190] The enzymes used had specific characteristics; no adverse effect on dough rheology, on crumb structure or the volume of the resulting bread; low activity at room temperature; and a relatively high optimum temperature. Such an activity pattern guarantees limited protein breakdown at temperatures encountered during mixing and proofing and higher breakdown in the early stage of the baking process. The enzymes used showed similar effects on a range of baked products.

6.7 OTHER ENZYMES

Besides amylases, xylanases, lipases, oxidases and proteases, several other (classes of) enzymes have been investigated for their effects in bread making and several other types mentioned above have been reported to have beneficial effects on one or more characteristics of dough or bread. And indeed some of them have been commercialized.

6.7.1 Transglutaminase

In the preparation of yeast-leavened goods from weak wheat flour, the dough often has an unsatisfactory stability. Such a dough has not been able to retain the carbon dioxide gas that is formed during fermentation. For this reason, it is generally common to add oxidizing compounds to the flour to improve the resistance to stretching.[191] Because of the effort to reduce the addition of chemicals to foodstuffs and instead to use natural auxiliaries, like enzymes, the problem arose of improving the resistance of dough without the addition of inorganic chemicals. It has been found that transglutaminase improves the resistance of dough, particularly yeast dough from wheat flour, in a manner comparable to potassium bromate.[192,193] Transglutaminase (TGase; synonym: Protein-Glutamine-γ-Glutamyltransferase and Protein-glutamine:amine γ-glutamyl-transferase; **EC 2.3.2.13**) is an enzyme accessible from various sources that is widespread in the animal and plant kingdoms. It is known that transglutaminase has a cross-linking effect on proteins independent of the redox system of the dough, not involving the thiol groups and disulphide bonds in the dough. The basic reaction is shown in Fig. 6.12.

However, the effect on the rheological properties of dough is similar to oxidation and the effects are the result from an increased number of disulphide bonds. The need for TGase for improvement of the stretching properties depends in individual cases on the nature of the flour. TGase can be added in different ways in the preparation of baked goods. The enzyme preparation can be used together with the remaining components of a bread improver system, but it can also be mixed with the flour at the mill. This has the advantage that the dosage can be based on the properties of the flour, that is on its natural gluten properties. In this way, a flour of consistent baking properties can always be supplied to the baker. Thus, TGase is preferably added to weak wheat flours in order to strengthen gluten properties.

The effect of TGase on the dough can be visualized in an extensogram (see Table 6.5).

Addition of TGase leads to strongly increased dough resistance and a reduced extensibility. Combining TGase with a protease can overcome too strong effects of the coupling enzyme. The results of TGase addition on bread quality can be seen in Table 6.6.

In addition to the desired increase in the dough resistance, a reduction in extensibility is sometimes observed, which leads to an earlier breakage of the dough when testing extensibility. It has been found that this undesired side effect can be excluded by combining

Fig. 6.12 Schematic representation of the transglutaminase reaction.

Table 6.5 Effect of transglutaminase on dough properties.

Dough property	No enzyme	2000 U TGase kg^{-1}	2000 U TGase kg^{-1} + 0.3 g protease kg^{-1}
Water uptake (%)	54.1	54.2	54.1
Dough resistance (EE)	230	510	400
Extensibility (mm)	182	120	170
Ratio of resistance/extensibility	1.3	4.25	2.35
Energy (cm^2)	80	75	122

transglutaminase with a protease. Furthermore, the volume of baked goods is increased and their crumb properties are improved by the use of a protease. The dough prepared using TGase can be worked up into high-quality bakery products under conventional conditions; wheat bread, rolls and variety breads are representative of such products.

6.7.1.1 *Transglutaminase in gluten-free bread*

Celiac disease (CD) is a chronic enteropathy caused by the intake of gluten proteins from widely prevalent food sources, such as wheat, rye, barley and possibly even oats. The ingestion of gluten causes an inflammatory response resulting in the destruction of the villous structure of the small intestine.[194] Currently, the only effective treatment for CD is the strict lifelong renunciation of gluten-containing foods.[195] Because cereal products, and especially bread, are part of the basic diet in many countries, there is a high demand for gluten-free breads.

In view of the fact that gluten is the major structure forming protein in wheat bread and that gluten is responsible for the viscoelastic properties, it is technically a big challenge to produce high quality gluten-free bread. Various ingredients (gums, dairy powders, rice, sorghum, starches) mimicking the properties of gluten have been evaluated.[107,196–200] Also enzymes (amylases, xylanases, proteases) have been evaluated in the manufacturing of gluten-free products.[185,201]

One of the main problems associated with gluten-free bread is obtaining a good structure. Transglutaminase can be a tool to improve the structure of gluten-free breads. The quality of these products is significantly better with the formation of a stable protein network. When the enzyme is used in combination with the right protein substrates, for example milk proteins, or egg proteins, protein networks are formed, resulting in improved volume, crumb structure and overall quality.[202]

Table 6.6 Effect of transglutaminase on bread characteristics.

No.	Additive/kg flour	Dough quality	Volume (mL/10 rolls)	Crumb porosity
1	None	Soft	1850	Open
2	500 U TGase	Woolly, solid	2000	Fine
3	2000 U TGase	Woolly, short	1750	Dense, solid
4	2000 U TGase + 0.3 g protease	Woolly	2200	Fine
5	500 U TGase + 0.2 g ascorbic acid	Woolly, solid	2100	Fine
6	1000 U TGase + improver	Woolly, solid	2300	Fine

6.7.2 Endoglycosidases

Type II endoglycosidases are a category of hydrolases which are capable of cleaving specific internal glycosidic linkages found in glycoproteins. These endoglycosidases cleave all or part of the carbohydrate moiety from a glycoprotein, depending on the location of the reactive glycosidic linkage in the glycoprotein. Examples of Type II endoglycosidases include endo-β-*N*-acetylglucosaminidases (Endo-D, Endo-H (**EC 3.2.1.96**), Endo-L, Endo-CI, Endo-CII, Endo-F-Gal and Endo-F), endo-α-*N*-acetylgalactosaminidase, endo-β-*N*-galactosidases, peptide-*N*-(*N*-acetyl-β-glucosaminyl), aspergine amidase F (PNGaseF **EC 3.5.1.52**) and glycopeptide *N*-glycosidase (Peptide *N*-glycosidase **EC 3.2.2.18**).

Wheat gluten proteins are to a certain extent also glycosylated. Glycans were detected on both gliadin and glutenin polypeptides. Covalently aggregated LMW glutenins were shown to contain *N*-glycans with xylose, which demonstrated their sorting in the Golgi apparatus.[203]

Tests done with purified endoglycosidases revealed enhanced dough relaxation.[158] This was explained by assuming better gluten network formation after removing side chains from gluten proteins which could hinder the building-up of a gluten network.

Since endoglycosidases are an intrinsic side activity of almost all xylanases, cellulases, glucanases and pectinases, there was limited commercial interest in large-scale production of such enzymes. The trend seen in the last decade to produce more and more enzymes from GMOs and even producing protein-engineered enzymes has resulted in the availability of 'mono-component' xylanases; that is xylanases which are purer as a result of the manufacturing process. This trend may lead to increased interest in these endoglycosidase side activities, since it is clear from the above that observed differences in performance of various xylanases may very well be the result of the presence or absence of non-xylanase side activities.

6.7.3 Cellulases

For cellulose breakdown the combined action of several enzymes is required. Cellulase (endo-1,4-β-D-glucanase **EC 3.2.1.4**) is the most relevant one. Five general types of cellulases, based on the type of reaction catalyzed, can be identified.

(1) Endocellulase breaks internal bonds to disrupt the crystalline structure of cellulose and expose individual cellulose polysaccharide chains; (2) exo-cellulase cleaves 2–4 units from the ends of the exposed chains produced by endocellulase, resulting in the tetrasaccharides or disaccharide such as cellobiose. There are two main types of exo-cellulases (cellobiohydrolases (CBH); **EC 3.2.1.91**) – one type working progressively from the reducing end, and one type working progressively from the non-reducing end of cellulose; (3) cellobiase or β-glucosidase hydrolyzes the endocellulase product into individual monosaccharides; (4) oxidative cellulases that depolymerize cellulose by radical reactions, as for instance cellobiose dehydrogenase and (5) cellulose phosphorylases that depolymerize cellulose using phosphates instead of water. The breakdown of cellulose is schematically shown in Fig. 6.13.

Most fungal cellulases have a two-domain structure with one catalytic domain, and one cellulose-binding domain, that are connected by a flexible link. This structure is adapted for working on an insoluble substrate and it allows the enzyme to diffuse two-dimensionally on a surface in a caterpillar way. However, there are also cellulases (mostly endoglucanases) that lack a cellulose-binding domain. These enzymes might have a swelling function.

Wholemeal bread formulations differ from those of standard bread as the former contains a higher level of both water-soluble and insoluble fibre ingredients. Soluble fibres consist of water extractable (WE) arabinoxylan, β-glucans and gums, whereas insoluble fibre is

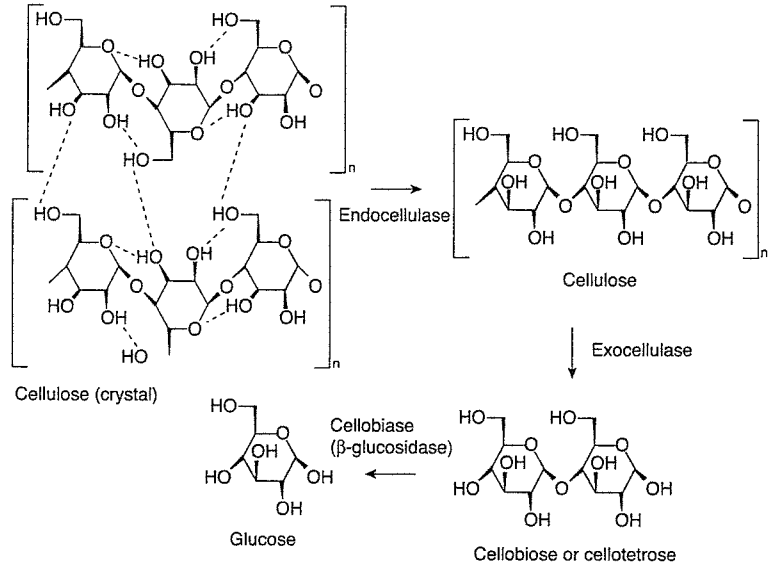


Fig. 6.13 Cellulose breakdown by endo- and exo-cellulases and cellobiase.

made up of lignin, cellulose and hemicellulose or water unextractable (WU) arabinoxylans. Wholemeal flour contains the entire seed of a plant and thus antioxidants, vitamins and fibres, which are beneficial to human health. Refined flour, used for standard white bread, has normally had the bran and the germ removed leaving only the endosperm.

The high fibre content of wholemeal bread may influence the dough consistency and weaken the gluten structure. This will ultimately lead to high water absorption, loss of extensibility and reduced fermentation tolerance. The final bread may show a reduced loaf volume, increased crumb firmness and a less pleasant taste and flavour. In addition, the crumb will be less white compared to normal bread.

Baking enzymes are therefore added to wholemeal formulations to improve dough and bread characteristics. In addition to the above-mentioned enzymes, cellulases are slowly finding their way into some bakery applications. The addition of cellulase enhances crumb structure, resulting in a more regular and fine crumb and, consequently, improved whiteness.[204]

The positive effect is due to break down of cellulose fibrils enabling better gluten development and improved proofing stability, increased dough tolerance and enhanced volume, texture and appeal in wholegrain bread types.

One of the relevant enzyme activities present in normal cellulase preparations is a CBH with bread improving activity.[205] The most important effects are an improved bread volume and an improved crumb structure, which are not accompanied by bad dough-handling properties due to stickiness of the dough (Table 6.7).

Like most sectors of the food industry, the bakery market has been presented with the opportunity of rising demand for healthy 'wellness' products promising nutritional benefits. Since consumption of whole grains has been linked to improvements in cardiovascular health as well as reducing the risk of certain cancers and lowering blood pressure, this move towards

Table 6.7 Effect of cellobiohydrolase on dough characteristics and loaf volume.

Addition	Dosage (ppm)	Dough quality	Volume (%)	Crumb structure
Reference		0	100	0
Xylanase	50	0	107	+
Xylanase	100	+	112	+
Cellobiohydrolase (CBH)	5	+	107	+
Cellobiohydrolase (CBH)	10	+++	110	++
Xylanase + CBH	50 + 5	+++	112	++
Xylanase + CBH	50 + 10	+++	115	+++

health-boosting foods has caused sales of wholegrain and high-fibre breads to rise rapidly in recent years with all major manufacturers offering wholegrain versions of their core bread brands to keep up with this demand.

6.7.4 Mannanases

β-Mannanase (1,4-β-D-mannan mannanohydrolase; **EC 3.2.1.78**) catalyzes the hydrolysis of β-1,4 mannose linkages of the backbone of β-mannans. These polysaccharides are found in various seeds and beans where they play an important role in the mechanical resistance and the swelling that occurs during germination. Mannans are also a major component of the hemicellulose fraction in soft woods.

There are several other types of enzymes that participate in the complete decomposition and conversion of the mannan, for example exo-β mannanase (1,4-β-D-mannan mannohydrolase; **EC 3.2.1.xx**–unassigned), exomannobiohydrolase (1,4-β-D-mannan mannobiohydrolase; **EC 3.2.1.100**) and β-mannosidase (**EC 3.2.1.25**).

Galactomannans and galactoglucomannans form a second group of hemicellulolytic structures present in plant cell walls. They are the major hemicellulose fraction of gymnosperms, in which they represent 12–15% of the cell wall biomass. Galactomannans are most commonly found in the family of Leguminoseae, in which they represent up to 38% of seed dry weight, but have also been identified in species of other plants such as Ebenaceae and Palmae. They consist of a backbone of β-1,4-linked D-mannose residues, which can be substituted by D-galactose residues via an α-1,6-linkage (Figs 6.14(a) and (b)). Depending on the source of the polysaccharide, mannose/galactose ratios can vary from 1.0 to 5.3.

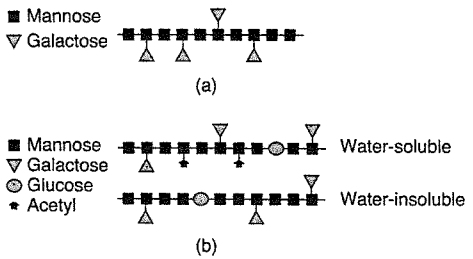


Fig. 6.14 (a) Schematic representation of the galactomannan structure. (b) Schematic representation of the galactoglucomannan structure.

Table 6.8 Effects of β-mannanase and guar on the anti-staling properties of bread.

Improver	Xylanase	Mannanase	Guar	Freshness (4 days)	Dough consistency
1	39.000	62	0	1	2
2	39.000	0	0.1%	2	3
3	39.000	62	0.1%	1	1

Galactoglucomannan is the major hemicellulolytic component of softwood. Two different structures can be identified within this group of polysaccharides. Both consist of a β-1,4-linked D-mannose backbone, which can be substituted by α-1,6-linked D-galactose. The galactoglucomannan backbone also contains β-1,4-linked D-glucose residues. Water-soluble galactoglucomannan has a higher galactose content than does water-insoluble galactoglucomannan and in addition contains acetyl residues attached to the main chain. Approximately, 20–30% of the backbone glucose and/or mannose residues are esterified with acetyl groups at C-2 or C-3.

Mannanases are useful in several industrial processes, such as the extraction of vegetable oil from leguminous seeds and the reduction of the viscosity of coffee extracts during manufacture of instant coffee. In the pulp and paper industry, mannanases can act synergistically with xylanases as biological pre-bleaching agents for soft wood pulp, allowing a significant reduction of environmental pollution compared with the use of chlorine-based chemical agents. However, in spite of these potentially interesting applications, the widespread use of mannanases is limited due to high production costs for the enzyme.

Mannanases are produced by microorganisms such as mould or yeast fungus as well as *Bacillus subtilis*, *Aeromonas*, *Enterococcus*, *Pseudomonas* and *Streptomyces*. Some higher plants or animals can produce mannanase. Microorganisms used for production of the mannanase are usually *Trichoderma* or *Aspergillus* sp. fungal strains.

Industrial applications of mannanases in bread making have been described.[76,206] The use of a β-mannanase in a baking improver in the preparation of baked goods improves the properties of dough and that of the final baked product. With regards to the dough, the tolerance can be improved, but also the dough flexibility, dough stickiness and general handling properties. With regards to the final baked product, mannanase retards staling and improves crumb structure.[206] Interesting is that the addition of a mannanase together with glucomannan and/or galactomannan such as guar of konjac gum, improves the properties of the dough and the baked products even further. This is shown in Tables 6.8 and 6.9.[206]

Table 6.9 Effect of specific enzymes on dough characteristics and shape.

	Dough		Shape with proofing time		Specific volume with proofing time	
Improver	Consistency	Stability	50′	65′	50′	65′
KBrO3	1	–	6	4	100	100
Xylanase (Xyl.)	4	+	7	5	104	103
Xyl. + glucose oxidase	2	+	7	5	106	105
Xyl. + β-mannanase	3	+	7	5	106	105
Xyl. + β-mannanase + glucose oxidase	2	++	7	6	107	107

6.8 CONCLUDING REMARKS

In the above, enzymes for baking have been discussed with regards to their classification, mechanism of action, effect and commercial availability. The word enzyme needs a little extra explanation. When talking about commercial enzyme preparations, these products are often referred to as, for example, lipase or xylanase or amylase. However, commercial preparations are hardly ever single enzyme preparations. In most cases commercial products are blends of enzymes. Besides the main enzyme activity, which often gives the preparation its name, a range of other enzymes can be present. These other enzymes are either present as natural side activities coming from the microorganism producing the main activity, or are deliberately added. This makes comparisons of commercial products very difficult, especially when looking only at the main activity. For a good comparison, it is absolutely necessary to test the products in the final application before a comparison is made based on activity, dose rate or product costs.

REFERENCES

1. Cura, J.A., Jansson, P.-E. and Krisman, C.R. (1995) Gelatinization of wheat starch as modified by xanthan gum, guar gum, and cellulose gum. *Starch* **47**, 207–209.
2. Atwell, W.A., Hood, L.F., Lineback, D.R., Varriano-Marston, E. and Zobel, H.F. (1988) Technology and methodology associated with basic starch phenomena. *Cereal Foods World* **33**, 306–311.
3. Schofield, J.D., Bottomley, R.C., Timms, M.F. and Booth, M.R. (1983) The effect of heat on wheat gluten and the involvement of Sulfhydryl-disulphide interchange reactions. *Journal of Cereal Science* **1**, 241–253.
4. Shewry, P. R. (1995) Plant storage proteins. *Biological Reviews* **70**, 375–426.
5. Wrigley, C.W. and Bietz, J.A. (1988) Proteins and amino acids. In: *Wheat Chemistry and Technology*, Vol. **1** (ed. Y. Pomeranz). AACC Inc., St. Paul, MN, pp. 159–275.
6. Finney, K.F. and Barmore, M.A. (1948) Loaf volume and protein content of hard winter and spring wheats. *Cereal Chemistry* **25**, 291–312.
7. Osborne, T.B. (1907) *The Proteins of the Wheat Kernel*. Publications of the Carnegie Institution Washington Judd and Detweiler, Washington, D.C.
8. Orth, R.A. and Bushuk, W. (1972) A comparative study of the proteins of wheats of diverse baking qualities. *Cereal Chemistry* **49**, 268–275.
9. Hoseney, R.C. (1994) *Principles of Cereal Science and Technology*, 2nd edn. Association of Cereal Chemists, Inc., St. Paul, MN, pp. 81–101, 229–273.
10. Kahn, K. (2006) Shelf life of bakery products. In: *Bakery Products, Science and Technology* (ed. Y.H. Hui). Blackwell Publishing Company, Ames, IA.
11. Belton, P.S. (1999) On the elasticity of wheat gluten. *Journal of Cereal Science* **29**, 103–107.
12. Dimler, R.J. (1963) Gluten, the key to wheat's utility. *Baker's Digest* **37**, 52–57.
13. Moonen, J.H.E., Scheepstra, A. and Graveland, A. (1986) Use of the SDS sedimentation test and SDS polyacrylamide gel electrophoresis for screening breeder's samples of wheat for bread making quality. *Euphytica* **31**, 677–690.
14. Payne, P.I., Nightingale, M.A., Krattiger, A.F. and Holt, L.M. (1987) The relationship between HMW glutenin subunit composition and the bread making quality of British-grown wheat varieties. *Journal of the Science of Food and Agriculture* **40**, 51–65.
15. Weegels, P.L., Hamer, R.J. and Schofield, J.D. (1996) Depolymerization and re-polymerization of wheat glutenin during dough processing I. Relationships between glutenin macropolymer content and quality parameters. *Journal of Cereal Science* **23**, 103–111.
16. Weegels, P.L., Hamer, R.J. and Schofield, J.D. (1997) Depolymerization and re-polymerization of wheat glutenin during dough processing II. Changes in composition. *Journal of Cereal Science* **25**, 155–163.

17. Bekes, F., Gras, P.W. and Gupta, R.B. (1994) Mixing properties as measurement of reversible reduction and oxidation of dough. *Cereal Chemistry* **71**, 44–50.
18. Jelaca, S.L. and Hlynca, I. (1971) Water binding capacity of wheat flour crude pentosans and their relation to mixing characteristics of dough. *Cereal Chemistry* **48**, 211–222.
19. Rattan, O., Izydorczyk, M. and Biliaderis, C.G. (1994) Structure and rheological behaviour of arabinoxylans from Canadian bread wheat flours. *Food Science and Technology* **27**, 350–355.
20. Sosulski, F., Krygier, K. and Hogge, L. (1982) Free, esterified and insoluble-bound phenolic acid. III. Composition of phenolic acids in cereal and potato flours. *Journal of Agricultural and Food Chemistry* **30**, 337–340.
21. Delcour, J.A., Vanhamel, S. and Hoseney, R.C. (1991) Physicochemical and functional properties of rye non-starch polysaccharides. II. Impact of a fraction containing water soluble pentosans and proteins on gluten-starch loaf volumes. *Cereal Chemistry* **68**, 72–76.
22. Michniewicz, J., Biliaderis, C.G. and Bushuk, W. (1991) Effects of added pentosans on physical and technological characteristics of dough and gluten. *Cereal Chemistry* **68**, 252–258.
23. Labat, E., Morel, M.H. and Rouau, X. (2000) Effects of laccase and ferulic acid on wheat flour dough. *Cereal Chemistry* **77**, 823–828.
24. Wang, M.W. (2003) Effect of pentosans on gluten formation and properties. PhD Thesis, Wageningen University.
25. Patil, S.K., Tsen, C.C. and Lineback, D.R. (1975) Water soluble pentosans of wheat flour. I. Viscosity properties and molecular weights estimated by gel filtration. *Cereal Chemistry* **52**, 44–56.
26. Geissmann, T. and Neukom, H. (1973) On the composition of the water soluble wheat flour pentosans and their oxidative gelation. *Lebensmittel-Wissenschaft und Technologie* **6**, 59–62.
27. Hoseney, R.C. and Faubion, J.M. (1981) A mechanism for the oxidative gelation of wheat flour water-soluble pentosans. *Cereal Chemistry* **58**, 421–424.
28. Izydorczyk, M.S., Biliaderis, C.G. and Bushuk, W. (1990) Physical properties of water-soluble pentosans from wheat. *Journal of Cereal Science* **11**, 153–169.
29. Rouau, X., El-Hayek, M.L. and Moreau, D. (1994) Effect of an enzyme preparation containing pentosanases on the bread making quality of flours in relation to changes in pentosans properties. *Journal of Cereal Science* **19**, 259–272.
30. Jelaca, S.L. and Hlynca, I. (1972) Effect of wheat flour pentosans in dough, gluten and bread. *Cereal Chemistry* **49**, 489–495.
31. Kim, S.K. and D'Appolonia, B.L. (1977a) Bread staling studies I: effect of protein content on staling rate and bread crumb pasting properties. *Cereal Chemistry* **54**, 207–215.
32. Kim, S.K. and D'Appolonia, B.L. (1977b) Bread staling studies II: effect of protein content and storage temperature on the role of starch. *Cereal Chemistry* **54**, 216–224.
33. Weegels, P.L., Marseille, J.P. and Hamer, R.J. (1992) Enzymes as processing aid in the separation of wheat flour into starch and gluten. *Starch* **44**, 44–48.
34. Jackson G.M. and Hoseney, R.C. (1986) Effect of endogenous phenolic acids on the mixing properties of wheat flour doughs. *Journal of Cereal Science* **4**, 79–85.
35. Neukom, H. and Markwalder, H.U. (1978) Oxidative gelation of wheat flour pentosans: a new way of cross-linking polymers. *Cereal Foods World* **23**, 374–376.
36. Figueroa-Espinoza, M.C., Morel, M.-H. and Rouau, X. (1998) Oxidative cross-linking of pentosans by a fungal laccase and horse radish peroxidase. Mechanism of linkage between feruloylated arabinoxylans. *Cereal Chemistry* **75**, 259–265.
37. Moore, A.M., Martinez-Munoz, I. and Bushuk, W. (1990) Factors affecting the oxidative elation of wheat water-solubles. *Cereal Chemistry* **67**, 81–84.
38. Oudgenoeg, G., Dirksen, E., Ingemann, S., Hilhorst, S.R., Gruppen, H., Boeriu, C.G., Piersma, S.R., Berkel, W.J.H., Laane, C. and Voragen, A.G. (2002) Horseradish peroxidase catalyzed oligomerization of ferulic acid on a template of a tyrosine containing tripeptide. *Journal of Biological Chemistry* **277**, 21332–21340.
39. Oudgenoeg, G., Hilhorst, R., Piersma, S.R., Boeriu, C.G., Gruppen, H., Hessing, M., Voragen, A.G. and Laane, C. (2001) Peroxidase mediated cross-linking of a tyrosine containing peptide with ferulic acid. *Journal of Agricultural and Food Chemistry* **49**, 2503–2510.
40. Vinkx, C.J.A., Van Nieuwenhove, C.G. and Delcour, J.A. (1991) Physicochemical and functional properties of rye non-starch polysaccharides III. Oxidative gelation of a fraction containing water-soluble pentosans and proteins. *Cereal Chemistry* **68**, 617–622.
41. Pomeranz, Y. and Chung, O.K. (1978) Interactions of lipids with proteins and carbohydrates in bread making. *Journal of the American Oil Chemists' Society* **55**, 285–289.

42. MacMurray, T.A. and Morrison, W.R. (1970) Composition of wheat flour lipids. *Journal of the Science of Food and Agriculture* **21**, 520.
43. Eliasson, A.-C. and Larsson, K. (1993) *Cereals in Bread Making. A Molecular Colloidal Approach*. Marcel Dekker Inc., New York, p. 376.
44. Morrison, W.R., Law, R.V. and Snape, C.E. (1993) Evidence for the inclusion complexes of lipids with V-amylose in maize, rice and oat starches. *Journal of Cereal Science* **18**, 107–109.
45. Chung, O.K. (1986) Lipid protein interactions in wheat flour, dough, gluten and protein fractions. *Cereal Foods World* **31**, 242–246.
46. Larsson, K. (1983) *Lipids in Cereal Technology*. Academic Press, London, pp. 237–251.
47. Gan, Z., Ellis, P.R. and Schofield, J.D. (1995) Amylose is not strictly linear. *Journal of Cereal Science* **21**, 215–230.
48. Kragh, K. (2002) Amylases in baking. In: *Recent Advances in Enzymes in Grain Processing* (eds C.M. Courtin, W.S. Veraverbeke and J.A. Delcour). Catholic University Leuven, Belgium, pp. 221–227.
49. Qi Si, J. (1996) New enzymes for the baking industry. *Food Technology Europe* **3**, 60–64.
50. Martinez-Anaya, M.A., Devessa, A., Andreu, P., Escriva, C. and Collar, C. (1999) Effects of the combination of starters and enzymes in regulating bread quality and shelf life. *Food Science and Technology International* **5**, 263–273.
51. van Duijnhoven, A.M. (2008) Personal Communication.
52. AFMB-CNRS-Universités Aix-Marseille I & II (1999) http://afmb.cnrs-mrs.fr/CAZY/
53. Qi Si, J. and Simmonsen, R. (1994) Functional mechanism of some microbial amylases antistaling effect. In: *Proceedings of the International Symposium on New Approaches in the Production of Food Stuffs and Intermediate Products from Cereal Grains and Oilseeds*. Beijing, China.
54. Pritchard, P. (1986) Studies on the bread-improving mechanism of fungal alpha amylase. *FMBRA Bulletin* **5**, 208–211.
55. van Dam, H.W. and Hille, J.D.R. (1992) Yeast and enzymes in bread making. *Cereal Foods World* **37**, 245.
56. Svensson, B. (1995) Protein engineering in the α-amylase family: catalytic mechanism, substrate specificity. *Plant Molecular Biology* **25**, 141–157.
57. Bowles, L.K. (1996) Amylolytic enzymes. In: *Baked Goods Freshness, Technology, Evaluation and Inhibition of Staling* (eds R.E. Hebeda and H.F. Zobel). Marcel Dekker Inc., New York, pp. 105–129.
58. Kulp, K. and Ponte, J.G. (1981) Staling of white pan bread: fundamental causes. *CRC Critical Reviews Food Science and Nutrition* **15**, 1–48.
59. Schoch, T.J. and French, D. (1947) Studies on bread staling. I: the role of starch. *Cereal Chemistry* **24**, 231–249.
60. MacRritchie, F. (1980) *Advances in Cereal Science and Technology III* (ed. Y. Pomeranz). American Association of Cereal Chemists, St. Paul, MN, Chapter 7.
61. Dragsdorf, R.D. and Varriano-Marston, E. (1980) Bread staling: X-ray diffraction studies on bread supplemented with α-amylases from different sources. *Cereal Chemistry* **57**, 310–314.
62. Martin, M.L. and Hoseney, R.C. (1991b) A mechanism of bread firming. II: role of starch hydrolysing enzymes. *Cereal Chemistry* **68**, 503–507.
63. Martin, M.L., Zeleznak, K.J., Hoseney, R.C. (1991a) A mechanism of bread firming. I: role of starch swelling. *Cereal Chemistry* **68**, 498–503.
64. Akers, A.A. and Hoseney, R.C. (1994) Water soluble dextrins from α-amylase treated bread and their relationship to bread firming. *Cereal Chemistry* **71**, 223–226.
65. Qi Si, J. (1994) Novo Nordisk applicant: use of laccase in baking. International Patent Application, WO 9428728.
66. Gerrard, J.A., Every, D., Sutton, K.H. and Gilpin, M.J. (1997) The role of maltodextrins in the staling of bread. *Journal of Cereal Science* **26**, 201–209.
67. Qi Si, J. and Lustenberger, C. (2002) Enzymes for bread, pasta and noodle products. In: *Enzymes in Food Technology* (eds R.J. Whitehurst and B.A. Law). CRC Press, Sheffield, pp. 19–57.
68. Anon. (2003) E.S.L. revolution/evolution. *Milling and Baking News*, March.
69. van Duijnhoven, A.M., Sturkenboom, M. and De Levita, P. (2005) Pan release agent. WO2005/094599.
70. Schieberle, P. (1990) The role of free amino acids present in yeast as precursors of the odorants 2-acetyl-pyrroline and 2-acetotetrahydropyridine in wheat bread crust. *Zeitschrift fur Lebensmittel Untersuchung und Forschung* **191**, 206–209.
71. Martinez-Anaya, M. and Jimenez, T. (1997) Functionality of enzymes that hydrolyses starch and non-starch polysaccharides in bread making. *Zeitschrift fur Lebensmittel Untersuchung und Forschung* **205**, 209–214.

72. Qi Si, J. (1988) Novamyl, a true antistaling enzyme. *Proceedings of IATA Meeting*, Valencia, Spain.
73. Anon. (2001) Novamyl is a maltogenic α-amylase. *Novozymes product sheet B547*.
74. Christophersen, C., Otzen, D.E., Norman, B.E., Christensen, S. and Schaefer, T. (1998) Enzymatic characterization of maltogenic alpha amylase, a thermostable α-amylase. *Starch* **50**, 39–45.
75. McCleary, B.V., Gibson, T.S., Allen, H. and Gams, T.C. (1986) Enzymic hydrolysis and industrial importance of barley glucans and wheat flour pentosans. *Starch* **38**, 433–437.
76. Maat, J., Roza, M., Verbakel, J., Stam, H., Santos da Silva, M.J., Bosse, M., Egmond, M.R., Hagemans, M.L.D., van Gorcom, R.F.M., Hessing, J.G.M and van den Hondel, C.A.M.J.J. (1992) Xylanases and their application in bakery. In: *Xylans and Xylanases* (eds G.B.J. Visser, M.A. Kusters-van Someren and A.G.J. Voragen). Elsevier Science Publishers, Amsterdam, Netherlands, pp. 349–360.
77. Gruppen, H., Kormelink, F.J.M. and Voragen, A.G.J. (1993) Water-unextractable cell wall material from wheat flour. III. A structural model for arabinoxylans. *Journal of Cereal Science* **19**, 11–18.
78. Rouau, X. and Moreau, D. (1993) Modification of some physicochemical properties of wheat flour pentosans by an enzyme complex recommended for baking. *Cereal Chemistry* **70**, 626–632.
79. Hamer, R.J. and Lichtendonk, W.J. (1987) Structure-function studies on gluten proteins. In: *Proceedings of the 3rd International Workshop on Gluten Proteins* (eds R. Lasztity and F. Bekezs). Budapest, Hungary, p. 227.
80. Courtin, C.M., Roelants, A. and Delcour, J.A. (2001) The use of two endoxylanases with different substrate selectivity provides insight into the role of endoxylanases in bread making. *Journal of Agricultural and Food Chemistry* **47**, 1870–1877.
81. Hamer, R.J. (1991) Enzymes in the baking industry. In: *Enzymes in Food Processing* (eds G.A. Tucker and L.F.J. Woods). Blackie, Glasgow, pp. 168–193.
82. Weegels, P.L. and Hamer, R.J. (1989) Predicting the baking quality of gluten. *Cereal Foods World* **34**, 210–212.
83. Wong, K.K.Y., Tan, L.U.L. and Saddler, J.N. (1988) Multiplicity of β-1,4-xylanase in microorganisms: functions and applications. *Microbiological Reviews* **52**, 305.
84. Jeffries, T.W. (1996) Biochemistry and genetics of microbial xylanases. *Current Opinion in Biotechnology* **7**, 337–342.
85. Biely, P., Vranska, M., Tenkanen, M. and Kluepfel, D. (1997) Endo-beta-1,4-xylanase families: differences in catalytic properties. *Journal of Biotechnology* **57**, 151–166.
86. Subramaniyan, S. (2000) Studies on the production of bacterial xylanases. PhD Thesis, Cochin University of Science and Technology, Kerala, India.
87. Bailey, M.J., Buchert, J. and Viikari, L. (1993) Effect of pH on production of xylanase by Trichoderma reesei on xylan and cellulose based media. *Applied Microbiology and Biotechnology* **40**, 224.
88. Subramaniyan, S. and Prema, P. (2003) Biotechnology of microbial xylanases: enzymology, molecular biology and application. *Critical Reviews in Biotechnology* **22**, 33–64.
89. Henrissat, B. and Bairoch, A. (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *The Biochemical Journal* **293**, 781–788.
90. Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. In: *Recent Advances in Carbohydrate Bioengineering* (eds H.J. Gibert, G. Davies, B. Henrissat and B. Svensson). The Royal Society of Chemistry, Cambridge, pp. 3–12. http://afmb.cnrs-mrs.fr/CAZY
91. Collins, T., Gerday, C. and Feller, G. (2005) Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiology Reviews* **29** (1), 3–23.
92. Collins, T., Meewis, M.A., Stals, I., Claeyssens, M., Feller, G. and Gerday, C. (2002) A novel family 8 xylanase, function and physicochemical characterization. *Journal of Biological Chemistry* **277**, 35133–35139.
93. Collins T., Gerday, C. and Feller, G. (2003) Xylanases, xylanase families and extremophilic xylanases. *Journal of Molecular Biology* **328**, 419.
94. van Petegem, F., Collins, T., Meewis, M.A., Feller, G. and van Beeumen, J. (2003) The structure of a cold adapted family 8 xylanase at 1.3 Å resolution, structural adaptations to cold and investigation of the active site. *Journal of Biological Chemistry* **278**, 7531–7539.
95. Courtin, C.M., Roelants, A. and Delcour, J.A. (1999) Fractionation reconstitution experiments provide insight into the role of endoxylanases in bread making. *Journal of Agricultural and Food Chemistry* **47**, 1870–1877.

96. Moers, K., Courtin, C.M., Brijs, K. and Delcour, J.A. (2002) A screening method for endo-beta-1,4 xylanase substrate selectivity. In: *Recent Advances in Enzymes in Grain Processing* (eds C.M. Courtin, W.S. Veraverbeke and J.A. Delcour). Catholic University Leuven, Belgium.
97. Guy, R.C.E. and Sahi, S.S. (2002) Comparison of effects of xylanases with fungal amylases in five flour types. In: *Recent Advances in Enzymes in Grain Processing* (eds C.M. Courtin, W.S. Veraverbeke and J.A. Delcour). Laboratory of Food Chemistry, Catholic University Leuven, Belgium.
98. Sprössler, B.G. (1997) Xylanases in baking. In: *1st European Symposium on Enzymes in Grain Processing* (eds S.A.G.F. Angelino, R.J. Hamer, W. van Hartingsveldt, F. Heidekamp and J.P. van der Lugt). TNO Food & Nutrition, Zeist, Netherlands, pp. 177–187.
99. Mathewson, P.R. (1998) Common enzyme reactions. *Cereal Foods World* **43**, 798–803.
100. Verger, R. and De Haas, G.H. (1973) Enzyme reactions in a membrane model. 1. A new technique to study enzyme reactions in monolayers. *Chemistry and Physics of Lipids* **10**(2), 127–136.
101. Brockerhof, H. and Jensen, G. (1974) In: *Lipolytic Enzymes* (eds J. Caro, M. Boudouard, J. Bonicel, A. Guidom and J. Desnuelle). Academic Press, New York.
102. Primo-Martin, C., Hamer, R.J. and de Jongh, H.H.J. (2006) Surface layer properties of dough liquor components: are they key parameters in gas retention in bread dough? *Food Biophysics* **1**, 83–93.
103. van Vliet, T., Janssen, A.M., Bloksma, A.H. and Walstra, J. (1992) Strain hardening of dough as a requirement for gas retention. *Journal of Texture Studies* **23**, 439.
104. Gan, Z., Angold, R.E., Pomeranz, Y., Shogren, M.D. and Finney, K.F. (1990) Response to shortening addition and lipid removal in flours that vary in bread making quality. *Journal of Cereal Science* **12**, 15–24.
105. Macritchie, F. and Gras, P.W. (1973) The role of flour lipids in baking. *Cereal Chemistry* **50**, 292–297.
106. Lundkvist, H., Arskog, P.B., Erlandsen, L., Ipsen, R. and Wilde, P (2007) Interfacial properties of dough liquor from lipase modified dough. In: *AACC C&E Spring Meeting*. Montpellier, France, 2–4 May.
107. Olesen, T. and Qi Si, J. (1994) Use of lipase in baking. International Patent Application, WO 94/0403035.
108. Qi Si, J. (1997) Synergistic effects of enzymes for bread making. *Cereal Foods World* **42**, 802–807.
109. Christiansen, L. (2006) Novozymes A/S (2001). Lipopan^{TM} F BG Product Sheet.
110. Castello, P., Baret, J.L., Potus, J. and Nicolas, J. (2000) Technological and biochemical effects of exogenous lipases in bread making. In: *2nd European Symposium on Enzymes in Grain Processing* (eds T. Simoinen and M. Tenkanen). Espoo, Finland, pp. 193–200.
111. Castello, P., Jollet, S., Potus, J. and Nicolas, J. (1998) Effects of exogenous lipase on dough lipids during mixing of wheat flours. *Cereal Chemistry* **75**, 595–601.
112. Johnson, R.H. and Welch, E.A. (1968) Baked goods dough and method. US Patent Application, WO 3,368,903.
113. Poulsen, C. and Borch Søe, J. (1997) Effect and functionality of lipases in dough and bread. In: *1st Symposium of Enzymes in Grain Processing* (eds S.A.G.F. Angelino, R.J. Hamer, W. van Hartingsveldt, F. Heidekamp and J.P. van der Lugt). TNO Nutrition and Food Research Institute, Zeist, Netherlands, pp. 204–214.
114. Xu, F. (2005) Application of oxidoreductases: recent progress. *Industrial Biotechnology* **1**, 38–50.
115. Munro, A.W., Taylor, P. and Wilkinshaw, M.D. (2000) Structures of redox enzymes. *Current Opinion in Biotechnology* **11**, 369–376.
116. Dong, W. and Hoseney, R.C. (1995) Effects of certain bread making oxidants and reducing agents on dough rheological properties. *Cereal Chemistry* **72**, 58–64.
117. Bloksma, A.H. (1972) The relation between thiol disulfide contents of dough and its rheological properties. *Cereal Chemistry* **49**, 104–117.
118. Jorgensen, H. (1939) Further investigations into the nature of the action of bromate and ascorbic acid on the baking strength of wheat flour. *Cereal Chemistry* **16**, 51–60.
119. Bloksma, A.H. and Bushuk, W. (1988) Rheology and chemistry of dough. In: *Wheat Chemistry and Technology*, Vols. 1 and 2 (ed. Y. Pomeranz). AACC, St. Paul, MN.
120. Oort, M.G. van (1996) Oxidases in baking. *International Food Ingredients* **4**, 42–47.
121. Nicolas, J. and Potus, J. (2000) Interactions between lipoxygenase and other oxidoreductases in baking. In: *2nd European Symposium on Enzymes in Grain Processing* (eds T. Simoinen and M. Tenkanen). VTT Symposium, Espoo, Finland, pp. 103–120.
122. Wieser, H. (2003) The use of redox agents. In: *Bread Making: Improving Quality* (ed. S.P. Cauvain). Woodland Publishing Ltd., Cambridge, pp. 424–446.

123. Vermullapalli, V. and Hoseney, R.C. (1980) Glucose oxidase in bread making systems. *Cereal Chemistry* **75**, 439–442.
124. Primo-Martin, C., Valera, R. and Martinez-Anaya, M.A. (2003) Effect of pentosanase and oxidase on the characteristics of dough and the glutenin macropolymer (GMP). *Journal of Agricultural and Food Chemistry* **51**, 4673.
125. Mitani, M., Maeda, T. and Morita, N. (2003) Effects of various kinds of enzymes on dough properties and bread qualities. In: *Recent Advances in Enzymes in Grain Processing* (eds C.M. Courtin, W.S. Veraverbeke and J.A. Delcour). Laboratory of Food Chemistry, Leuven, Belgium, pp. 295–302.
126. Rosell, C.M., Wang, J., Aja, S., Bean, S. and Lockhart, G. (2003) Wheat flour proteins as affected by transglutaminase and glucose-oxidase. *Cereal Chemistry* **80**, 52–55.
127. Matheis, G. and Whitaker, J.R. (1987) A review: enzymatic cross-linking of food proteins applicable to foods. *Journal of Food Biochemistry* **11**, 309–327.
128. Haas, L.W. (1934) Bleaching agent and process of preparing bleached bread dough. United States Patent 1957334.
129. Hoseney, R.C., Rao, H., Faubion, J. and Sighu, J.S. (1980) Mixograph studies. IV. The mechanism by which lipoxygenase increases mixing tolerance. *Cereal Chemistry* **57**, 163–166.
130. Addo, K., Burton, D., Stuart, M.R., Burton, H.R. and Hildebrand, D.F. (1993) Soybean flour lipoxygenase isozyme mutant effects on bread dough volatiles. *Journal of Food Science* **58**, 583–585.
131. Delcros, J.F., Rakotozafy, L., Boussard, A., Davidou, S., Porte, C., Potus, J. and Nicolas, J. (1998) Effect of mixing conditions on the behaviour of lipoxygenase, peroxidase and catalase in wheat flour dough. *Cereal Chemistry* **75**, 85–93.
132. Junqueira, R.M., Rocha, F., Moreira, M.A. and Castro, I.A. (2007) Effect of proofing time and wheat flour strength on bleaching, sensory characteristics and volume of French breads with added soybean lipoxygenase. *Cereal Chemistry* **84**, 443–449.
133. Logan, J.L. and Learmonth, E.M. (1955) Gluten oxidizing capacity of soya. *Chemistry And Industry* **39**, 1220.
134. Grosch, W. (1986) Redox systems in dough. In: *The Chemistry and Physics of Baking: Materials, Processes and Products* (eds J.M. Blanshard, P.J. Frazier and T. Galliard). Royal Society of Chemistry, London, pp. 155–169.
135. Koch, R.B. (1956) Mechanisms of fat oxidation. *Bakers Digest* **30**, 48–53.
136. Dahle, L.K. and Sullivan, B. (1963) The oxidation of wheat flour. V. Effect of lipid peroxides and antioxidants. *Cereal Chemistry* **40**, 372–384.
137. Smith, D.E. and Andrews, J.S. (1957) The uptake of oxygen by flour dough. *Cereal Chemistry* **34**, 323–326.
138. Graveland, A. (1970) Enzymatic oxidations of linoleic acid and glycerol-1-monolinoleate in dough and flour water suspensions. *Journal of the American Oil Chemists' Society* **47**, 352–361.
139. Graveland, A. (1973) Analysis of lipoxygenase non volatile reaction products of linoleic acid in aqueous cereal suspensions by urea extraction and gas chromatography. *Lipids* **8**, 599–605.
140. Tait, S.P.C. and Galliard, T. (1988) Oxidation of linoleic acid in dough and aqueous suspensions of wholemeal flours: effect of storage. *Journal of Cereal Science* **8**, 55–67.
141. Levavasseur, L., Rakotozafy, L., Manceau, E., Louarne, L., Robert, H., Bartet, J-L., Potus, J. and Nicolas, J. (2006) Discrimination of wheat varieties by simultaneous measurements of oxygen consumption and consistency of flour during mixing. *Journal of the Science of Food and Agriculture* **86**, 1688–1698.
142. Tsen, C.C. and Hlynka, I. (1963) Flour lipids and oxidation of sulfhydryl groups in dough. *Cereal Chemistry* **40**, 145.
143. Tsen, C.C. (1965) The improving mechanism of ascorbic acid. *Cereal Chemistry* **42**, 86–96.
144. Bloksma, A.H. (1963) Oxidation by molecular oxygen of thiol groups in unleavened dough from normal and defatted wheat flours. *Journal of the Science of Food and Agriculture* **14**, 529–535.
145. Graveland, A. (1971) Modification of the Lipoxygenase reaction by wheat glutenin. PhD Thesis, University Utrecht, Utrecht.
146. Avram, E., Boussard, A., Potus, J. and Nicolas, J. (2003) Oxidation of glutathione by purified wheat and soybean Lipoxygenase in the presence of linoleic acid at various pH. In: *Recent Advances in Enzymes in Grain Processing* (eds C.M. Courtin, W.S. Veraverbeke and J.A. Delcour). Katholieke Universiteit, Leuven, Belgium.
147. Kiefer, R., Mattheis, G., Hofmann, G. and Belitz, H.D. (1982) *Zeitschrift Lebensmittel Untersuchung und Forschung*, **75**, 5–7.

148. Fukushige, H., Wang, C., Simpson, T., Gardner, H. and Hildebrand, D. (2005) Purification and identification of linoleic acid hydroperoxides generated by soybean seed lipoxygenases 2 and 3. *Journal of Agricultural and Food Chemistry* **53**, 5691–5694.
149. Nicolas, J., Richard-Forget, F., Goupy, P., Amiot, M.J. and Aubers, S. (1994) Enzymatic browning reactions in apples and apple products. *Critical Reviews in Food Science and Nutrition* **34**, 109–157.
150. Zawistowski, J., Biliaderis, C.G. and Eskin, N.A.M. (1991) Polyphenol oxidase. In: *Oxidative Enzymes in Foods* (eds D.S. Robinson and N.A.M. Eskin). Elsevier Applied Science Chemistry, London, pp. 217–273.
151. Kuninori, T., Nishiyama, J. and Matsumoto, H. (1976) Effect of mushroom extract on the physical properties of dough. *Cereal Chemistry* **53**, 420–428.
152. Takasaki, S. and Kawakishi, S. (1997) Formation of protein bound 3,4-dihydroxy-phenylalanine and 5-cysteinyl-3,4-dihydrophenylalanine as new cross-linkers in gluten. *Journal of Agricultural and Food Chemistry* **45**, 3472–3475.
153. Takasaki, S., Kawakishi, S., Murat, M. and Homma, S. (2001) Polymerization of gliadin mediated by mushroom Tyrosinase. *Lebensmittel Wissenschaft Technologie*, **34**, 507–512.
154. Tilley, K.A. (1999) Method of dough manufacture by monitoring and optimizing gluten protein linkages. US patent No 6.284.296.
155. Tilley, K.A., Benjamin, R.E., Bagorogoza, K.E., Okot-Kolber, B.M., Prakash, O. and Kwen, H. (2000) Tyrosine cross links: The molecular basis of gluten structure and function. *Journal of Agricultural and Food Chemistry* **49**(6), 2627–2632.
156. Hillhorst, R., Gruppen, H., Orsel, R., Laane, C., Schols, H.A. and Voragen, A.G.J. (2000) On the mechanism of action of peroxidase in wheat dough. In: *2nd European Symposium on Enzymes in Bread Making* (eds T. Simoinen and M. Tenkanen). Techn. Research Centre of Finland VTT, Espoo, Finland, pp. 127–132.
157. Reinikainen, T., Lantto, R., Niku-Pavoola, M-L. and Buchert, J. (2003) Enzymes for cross-linking of cereal polymers. In: *Recent Advances in Enzymes in Grain Processing* (eds C.M. Courtin, W.S. Veraverbeke and J.A. Delcour). Laboratory of Food Chemistry, Leuven, Belgium, pp. 91–99.
158. Oort, M.G. van, Hennink, H. and Moonen, H. (1997) Peroxidases in bread making. In: *First European Symposium on Enzymes in Grain Processing* (eds S.A.G.F. Angelino, R.J. Hamer, W. van Hartingsveldt, F. Heidekamp and J.P. van der Lugt). TNO Nutrition and Food Institute, Zeist, The Netherlands, pp. 195–203.
159. Dunnewind, B., Vliet, T. van and Orsel, R. (2002) Effect of oxidative enzymes on bulk rheological properties of wheat flour dough. *Journal of Cereal Science* **36**, 357–366.
160. Swaisgood, H.E. (1980) Oxygen activation by sulfhydryl oxidase and the enzyme's interaction with peroxidase. *Enzyme and Microbial Technology* **2**, 265–272.
161. Kaufman, S.P. and Fennema, O. (1987) Evaluation of sulfhydryloxidase as a strengthening agent for wheat flour dough. *Cereal Chemistry* **64**, 172–176.
162. Fok, J.J., Hille, J.D.R. and Ven, B. Van Der (1993) Yeast derivative to improve bread quality. European Patent 0588426A1.
163. Grosch, W. and Wieser, H. (1990) Redox reactions in wheat dough affected by ascorbic acid. *Journal of Cereal Science* **29**, 1–16.
164. Grant, D.R. and Sood, V.K. (1980) Studies in the role of ascorbic acid in chemical dough development. II. Partial purification and characterization of an enzyme oxidizing ascorbic acid in flour. *Cereal Chemistry* **57**, 46–49.
165. Every, D., Gilpin, M.J. and Larsen, N.G. (1995) Continuous spectroscopic assay and properties ascorbic acid oxidizing factors in wheat. *Journal of Cereal Science* **21**, 231–239.
166. Every, D., Gilpin, M.J. and Larsen, N.G. (1996) Ascorbic acid oxidase levels in wheat and relationship to baking quality. *Journal of Cereal Science* **23**, 145–151.
167. Cherkiatgumchai, P. and Grant, D.R. (1986) Enzymes that contribute to the oxidation of L-ascorbic acid in flour-water systems. *Cereal Chemistry* **63**, 197–200.
168. Grant, D.R. (1974) Studies in the role of ascorbic acid in chemical dough development. I. Reaction of ascorbic acid with flour water suspensions. *Cereal Chemistry* **51**, 584–592.
169. Sarwin, R., Laskawy, G. and Grosch, W. (1993) Changes in the levels of glutathione and cysteine during the mixing of dough with L-threo- and D-erythro-ascorbic acid. *Cereal Chemistry* **70**, 553–557.
170. Kanht, W.D., Murdy, V. and Grosch, W. (1975) Verfahren zur bestimmung der aktivität des glutathion-dehydrogenase (EC 1.8.5.1.). Vorkommen des enzymes in verschiedenen weizensorten. *Zeitschrift Lebensmittel Untersuchung und Forschung* **158**, 77–82.

171. Arnaut, F., De Meyer, K. and Van Haesendonck, I. (2005) Bakery products comprising carbohydrate oxidase. European Patent EP1516536.
172. McDonald, E.C. (1969) Proteolytic enzymes of wheat and their relation to baking quality. *Baker's Digest* **43**, 26–28, 30 and 72.
173. Mathewson, P.R. (2000) Enzymatic activity during bread baking. *Cereal Foods World* **45**, 98–101.
174. Lindahl, L. and Eliasson, A.-C. (1992) Influence of added enzymes on rheological properties of a wheat flour dough. *Cereal Chemistry* **69**, 542–546.
175. Stauffer, C.E. (1987) Proteases, peptidases and inhibitors. In: *Enzymes and Their Role in Cereal Technology* (eds J.E. Kruger, D. Lineback and C.E. Stauffer). AACC Inc., St. Paul, MN, pp. 166–169.
176. Stauffer, C.E. (1994) In: *The Science of Cookie and Cracker Production* (ed. H. Faridi). Chapman & Hall, New York, pp. 237–238.
177. Nightingale, M.J., Marchylo, B.A., Clear, R.M., Dexter, J.E. and Preston, K.R. (1999) Fusarium head blight: effect of fungal proteases on wheat storage proteins. *Cereal Chemistry* **76**, 150–158.
178. Rosell, C.M., Aja, S., Bean, S. and Lookhart, G. (2002) Effect of Aelia spp. and Eurygaster spp. damage on wheat proteins. *Cereal Chemistry* **79**, 801–805.
179. Thiele, C., Gänzle, M.G. and Vogel, R.F. (2002) Contribution of sourdough Lactobacilli, Yeast and cereal enzymes to the generation of amino acids in dough relevant for bread flavour. *Cereal Chemistry* **79**, 45–51.
180. Barbeau, W.E., Griffey, C.A. and Yan, Z. (2006) Evidence that minor sprout damage can lead to significant reductions in gluten strength of winter wheats. *Cereal Chemistry* **83**, 306–310.
181. Schieberle, P. (1996) Intense aroma compounds – useful tools to monitor the influence of processing and storage on bread aroma. *Advances in Food Science* **18**, 237–244.
182. Gobetti, M., Simonetti, M.S., Rossi, J., Cossignani, L., Corsetti, A. and Damiani, P. (1994) Free D- and L-amino acid evolution during sourdough fermentation and baking. *Journal of Food Science* **59**, 881–884.
183. Hansen, A., Lund, B. and Lewis, M.J. (1989) Flavour of sourdough rye bread crumb. *Lebensmittel Wissenschaft und Technologie* **22**, 141–144.
184. Loponen, J., Mikola, M., Sontag-Strohm, T. and Salovaara, H. (2002) Degradation of high molecular weight glutenin subunits during wheat sourdough fermentation. In: *Recent Advances in Enzymes in Grain Processing* (eds C.M. Courtin, W.S. Veraverbeke and J.A. Delcour). Lab. Food Chem, Kath. University Leuven, Leuven, Belgium, pp. 281–287.
185. Harada, O., Lysenko, E.D. and Preston, K. R. (2000) Effects of commercial hydrolytic enzyme additives on Canadian short process bread properties and processing characteristics. *Cereal Chemistry* **77**, 70–76.
186. Harada, O., Lysenko, E.D., Edwards, N.M. and Preston, K.R. (2005) Effects of commercial hydrolytic enzyme additives on Japanese style sponge and dough bread properties and processing characteristics. *Cereal Chemistry* **82**, 314–320.
187. Barrett, A.H., Marando, G., Leung, H. and Kaletunc, G. (2005) Effect of different enzymes on textural stability of shelf stable bread. *Cereal Chemistry* **82**, 152–157.
188. Linares, E., Larre, C., Lemeste, M. and Popineau, Y. (2000) Emulsifying and foaming properties of gluten hydrolysates with an increasing degree of hydrolysis: role of soluble and insoluble fractions. *Cereal Chemistry* **77**, 414–420.
189. Cluskey, J.E. (1959) Relation of the rigidity of flour, starch, and gluten gels to bread staling. *Cereal Chemistry* **36**, 236–246.
190. Arnout, F., Verte, F. and Vekemans, N. (2003) Method and composition for retarding staling of bakery products by adding a thermostable protease. European Patent EP1350432.
191. Pomeranz, Y. (1971) *Wheat Chemistry and Technology*. AACC, St. Paul, MN, p. 700.
192. Gottmann, K. and Sproessler, B. (1994) Baking products and intermediates. US Patent 5279839.
193. Gottmann, K. and Sproessler, B. (1995) Baking agent and process for the manufacture of dough and bakery products. European Patent EP 0492406.
194. Shan, L., Molberg, O., Parrot, I., Hausch, F., Filiz, F. and Gray, G.M. (2002) Structural basis for gluten intolerance in celiac sprue. *Science* **297**, 2275–2279.
195. Feighery, C. (1999) Fortnightly review: celiac disease. *British Medical Journal* **319**, 236–239.
196. Toufeili, I., Dahger, S., Shadarevian, S., Noureddine, A., Sarakbi, M. and Farran, M.T. (1994) Formulation of gluten-free pocket-type flat breads. Optimization of methylcellulose, gum Arabic and egg albumen levels by response surface methodology. *Cereal Chemistry* **71**, 594–601.

197. O'Brien, C.M., von Lehmden, S. and Arendt, E.K. (2002a) Development of gluten free pizzas. *Irish Journal of Agriculture and Food Research* **42**, 134–137.
198. O'Brien, C.M., Schober, T. and Arendt, E.K. (2002b) Evaluation of the effect of different ingredients on the rheological properties of gluten-free pizza dough. *AACC Annual Meeting*. Published online at http://www.scioc.org/aacc/meeting/2002/abstracts. AACC International, St. Paul, MN.
199. Gallagher, E., Gormley, T.R. and Arendt, E.K. (2003) Recent advances in the formulation of gluten-free cereal based products. *Trends in Food Science & Technology* **15**, 143–152.
200. Gallagher, E., Gormley, T.R. and Arendt, E.K. (2004) Crust and crumb characteristics of gluten-free breads. *Journal of Food Engineering* **56**, 153–161.
201. Guarda, A., Rosell, C.M., Benedito, C. and Galotto, M.J. (2003) Different hydrocolloids as bread improvers and antistaling agents. *Food Hydrocolloids* **18**, 241–247.
202. Moore, M.M., Heinbockel, M., Dockery, P., Ulmer, H.M. and Arendt, E.K. (2006) Network formation in gluten-free bread with application of transglutaminase. *Cereal Chemistry* **83**, 28–36.
203. Laurière, M. and Denery, S. (2005) Céréales et dérivés. In: *Méthodes D'analyses Immunochimiques Pour le Contrôle de Qualité Dans les IAA* (eds J. Daussant and P. Arbault). Lavoisier, Ed Tec et Doc, Paris, pp. 293–328.
204. Maat, J. and Roza, M. (1995) Cellulase bread improvers e.g., xylanase include an oxidase or peroxidase. European Patent EP0396162.
205. van Beckhoven, R.F.W.C. (2003) Bread improving composition. US Patent 6656513.
206. van Duijnhoven, A.M. (1993) Enzyme containing baking improver. European Patent EP0529712.

Available online at www.sciencedirect.com
ScienceDirect
Food and Chemical Toxicology 44 (2006) 1909–1915
www.elsevier.com/locate/foodchemtox

Investigation on possible allergenicity of 19 different commercial enzymes used in the food industry

Carsten Bindslev-Jensen [a,*], Per Stahl Skov [b], Erwin L. Roggen [c], Peter Hvass [c], Ditte Sidelmann Brinch [c]

[a] *Department of Dermatology and Allergy Center, Odense University Hospital, DK 5000 Odense C, Denmark*
[b] *RefLab, National University Hospital, Tagensvej 20, Dep. 7512, 2200 Copenhagen N, Denmark*
[c] *Novozymes A/S, Krogshoejvej 36, DK 2880 Bagsvaerd, Denmark*

Received 28 October 2005; accepted 9 June 2006

Abstract

The aim of the study was to investigate the safety to allergic patients of 19 commercially available and authority-approved enzymes used in the food industry. Enzymes produced by genetically modified organisms were included. Four hundred consecutive adult patients with a diagnosed allergy to inhalation allergens, food allergens, bee or wasp were included. All had at least one positive skin prick test to the above allergens.

Skin prick testing with the 19 enzymes was performed on the forearm and if positive (in 13 patients), in vitro histamine release from blood basophils were performed. Patients with positive results in skin prick test were subsequently reinvestigated with further purified enzymes and finally challenged orally with the enzymes in a double-blind, placebo-controlled protocol. Only one reaction to a placebo challenge was seen.

In some instances a positive skin prick test result or a positive histamine release was seen elicited by the enzymes, but since none of the patients were positive to any of the commercial enzymes in the subsequent oral challenges using exaggerated dosages of the enzymes compared to normal daily intake, the findings are without clinical relevance.

A wide variety of enzyme classes and origins was included in the study. Because there were no allergenic findings of clinical relevance it is concluded that ingestion of food enzymes in general is not considered to be a concern with regard to food allergy.

Keywords: Allergen; Food allergy; Enzyme; Genetically modified; Human; Skin prick test; Basophil histamine release

1. Introduction

Food safety is of major concern worldwide, and one aspect of food safety is avoiding allergic reactions associated with food consumption. Patients with food hypersensitivity always face the risk of developing allergic symptoms after unintentional intake of a non-tolerated food. Such unintentional intake may be due to eating in a restaurant setting, where product labelling is lacking, or be due to insufficient labelling of food compounds in the daily diet.

As a result of the growing awareness and concerns of food allergy, regulatory agencies worldwide are developing allergen evaluation schemes and implementing allergen labelling regulations.

One example of such evaluation scheme is the decision tree as described in the FAO/WHO 2001 report (Aalberse et al., 2001), providing guidance on the assessment of the allergenic potential of foods derived from biotechnology. Such assessment may be required to protect food allergic patients against potential new risks associated with the development of genetically modified foods, whereby

* Corresponding author.
E-mail address: Carsten.Bindslev-Jensen@ouh.fyns-amt.dk (C. Bindslev-Jensen).

0278-6915/S - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.fct.2006.06.012

introduction of foreign allergenic proteins from other foods into a food which the patient has previously tolerated may elicit allergic reactions in the allergic patient (Hansen et al., 2004; Bindslev-Jensen et al., 2003).

For the allergen labelling regulations being implemented, this generally means mandatory declaration of recognized allergenic substances contained in the final foods (EU: Council Directive 2000/13/EC on the labelling, presentation and advertising of foodstuffs as amended by 2003/89/EC of 10 November, 2003. US: 21CFR101.4 and 21CFR101.100. Additionally, the 'Food Allergen Labelling and Consumer Protection Act' (FALCPA), was adopted August 2, 2004, to be effective January 1, 2006. JP: Food Allergen Labelling Guidelines of 2001. AU: Australia New Zealand Food Standards Code of 2000). Such mandatory labelling typically concerns the eight major food allergens ("the Big Eight": Cow's milk; Hen's egg; Fish; Crustacean/shellfish; Tree nuts; Wheat; Peanuts; and Soybeans), but may also comprise substances, which can cause intolerance only e.g. lactose.

As a result of above developments, all products being produced using modern biotechnology are subject to increased requirements on documenting the safety towards the final consumer from regulatory authorities, including appropriate documentation that ingestion of the products is not a concern for food allergy.

Enzymes constitute a class of products being produced by modern biotechnology. They may be extracted from animal or vegetable sources or may be produced in bacteria, fungi or yeasts. Some enzymes are produced by micro-organisms derived from wild-type strains (non-GMM); others are produced by genetically modified micro-organisms (GMM). Enzymes found in nature have been used since ancient time in the production of foods and in the manufacture of commodities. All these processes relied on either enzymes produced by spontaneously growing microorganisms or enzymes present in added preparations such as calves' rumen or papaya fruit. Industrially produced, food grade enzymes are used as processing aids in the manufacturing of a wide variety of foods such as bread, beer, beverages, dairy products etc.

Enzymes are proteins. Like many other proteins they may have the potential to cause allergic responses. Investigations have demonstrated that workers producing enzymes may develop sensitization to the enzymes after inhalation exposure (Bernstein et al., 1994, 1999; Zober, 2002; Merget et al., 2001; Burstyn et al., 1998; Leser et al., 2001; Quirce et al., 2002; van Kampen et al., 2002). This is today minimized considerable due to better knowledge of safe handling of the enzymes e.g. by making them as free of dust as possible e.g. by developing tough encapsulated granulates. In contrast, no reports on sensitization to these enzyme products in the final commercial food after ingestion exist. This may be due to the difference in exposure pattern of the inhalation route compared to the digestive route, or it may be due to the fact that the enzymes most often are present in the final foods in low amounts and in inactive forms.

The aim of the present study was to investigate the safety to allergic patients of 19 commercially available and authority approved enzymes. The enzymes were selected to reflect a wide range of enzymatic activities as well as enzymes being produced by both non-GMM and GMM.

2. Materials and methods

Four hundred consecutive adult patients (276 female and 124 men, mean age 38 years) with diagnosed allergy to inhalation allergens, food allergens, allergens of bee or wasp, or drugs were included, after informed consent was obtained. All patients had a positive skin prick test (SPT) result towards at least one of the above allergens according to EAACI guidelines (Dreborg et al., 1987) and all were free of symptoms on the day(s) of testing. Patients with a history of severe allergic reaction, pregnant or lactating women and patients treated with drugs known to interfere with the result of skin prick testing were excluded. The sensitizations in the patients are presented in Table 1.

The enzymes investigated are presented in Table 2. The enzyme preparations applied in the SPT and histamine release (HR) testing described in this paper were test batches which all were mixtures of 3–5 separately recovered and fermented sub test batches to secure that the testing material were representative. The test batches were fermented and recovered according to the same procedures as are used for production of commercial enzyme preparations. A part from water and the enzyme protein itself the test batch also contained other soluble organic substances from the fermentation; mainly protein and carbohydrate components. All test batches were analyzed extensively for chemical and microbial content documenting that the test batches complied with the Food and Agriculture Organization/World Health Organization (FAO/WHO) Joint Expert Committee on Food Additives (JECFA) and Food Chemical Codex (FCC) recommended purity specifications for food grade enzymes, including analyses to show that the test batches did not contain the production strain (FCC, 2001; JECFA, 2004).

The test batches are usually used for all the toxicological investigations done on the enzyme preparations before registration. The same test batches are also used in Novozymes' Occupational Health Service (OHS) for skin prick testing when testing workers for possible occupational allergy

Table 1
The different sensitizations represented and their distribution in the 400 patients included in the study, STEP 1

Positive skin prick test	Female	Male
Birch	149	46
Grass	164	77
Mugworth	62	14
Horse	42	10
Dog	94	21
Cat	94	26
Dust mites	95	47
Moulds	13	4
Vespula vulgaris	9	9
Honey bee	1	3
Hen's egg	4	1
Cow's milk	6	3
Tree nuts	47	14
Penicillin	4	1
Fish	3	8
Latex	1	1
Others (rabbit, tomato, wheat, rye, banana, poultry, peanut, poppy seed, shrimp, crab, guinea pig, chlorohexidine)	5	4

The patients may have a number of different sensitizations.

Table 2
The 19 enzymes tested in the study

Enzyme no.	Enzyme type	PE	Production strain	GMM	Donor organism
1	Maltogenic amylase	–	*Bacillus subtilis*	Yes	*Bacillus* sp. 1
2	Protease	–	*Bacillus amyloliquefaciens*	No	NA
3	Decarboxylase	–	*Bacillus subtilis*	Yes	*Bacillus* sp. 2
4	Alpha-amylase	+	*Bacillus licheniformis*	Yes	*Bacillus* sp. 3
5	Alpha-amylase	+	*Bacillus licheniformis*	Yes	*Bacillus* sp. 1
6	Glucoamylase	–	*Aspergillus niger*	Yes	*Aspergillus* sp. 1
7	Alpha-amylase	–	*Bacillus amyloliquefaciens*	No	NA
8	Pectin lyase	–	*Aspergillus niger*	Yes	*Aspergillus* sp. 1
9	Glucoseoxidase	–	*Aspergillus niger*	No	NA
10	Lipase	+	*Aspergillus oryzae*	Yes	Thermomyces sp. 1
11	Lipase	–	*Aspergillus oryzae*	Yes	Fusarium sp.
12	Xylanase	–	*Aspergillus oryzae*	Yes	Thermomyces sp. 1
13	Pectinesterase	–	*Aspergillus oryzae*	Yes	*Aspergillus* sp. 2
14	Beta-glucanase	–	Humicola insolens	No	NA
15	Glucoseoxidase	–	*Aspergillus oryzae*	Yes	*Aspergillus* sp. 1
16	Laccase	–	*Aspergillus oryzae*	Yes	Myceliopthora sp.
17	Alpha-amylase	–	*Aspergillus oryzae*	No	NA
18	Alpha-amylase	–	*Bacillus licheniformis*	Yes	*Bacillus* sp. 3
19	Protease	–	*Bacillus licheniformis*	No	NA

PE: protein-engineered enzyme, GMM: gene modified microorganism, sp: species, NA: not applicable.

against enzyme preparations. A concentration of 100 μg protein per ml in the SPT has been used by OHS for several years and has also been used for this study. A too high protein concentration will cause many false-positive reactions due to irritation and a concentration of 100 μg protein per ml has been found to have a sensitivity of 100% but a reduced specificity (Bernstein et al., 1994, 1993).

3. Test protocol

After informed consent was obtained the patients were tested using SPT in single determination with all the enzymes using the test batches with a concentration of 100 μg protein per ml and prepared in 50% glycerol (STEP 1). Any positive reaction (wheal > 3 mm larger than the negative control) was repeated in duplicate and if still positive, titrated in 1:10 dilutions in duplicate until the reaction had disappeared. Histamine hydrochloride 10 mg/ml was used as the positive control.

In all patients with a positive initial reaction in skin prick test, blood was drawn for histamine release testing with the enzyme test batch(es) in question.

Histamine release from basophil leukocytes was performed as previously described (Scheurer et al., 2001). Samples of 25 μl heparinized blood was applied to glass fibre coated microtitre wells (HR-Test from RefLab, Copenhagen, Denmark) and incubated with 25 μl of a dose range of the enzymes for 60 min at 37 °C. Each enzyme was tested in 12 concentrations, each in duplicate, from 100 μg/ml to 0.1 ng/ml (dilution factor 1:3.5). All serial dilutions of enzyme were made in PIPES-buffer (RefLab, Denmark). During incubation of patient blood with enzyme, released histamine is adsorbed to the glass microfibre coated microtiter plates followed by a fluorometrically determination of released histamine. A release of 10 ng histamine/ml blood is significant corresponding to 3 × standard deviation of background fluorescence. Unspecific enzyme induced histamine release was examined by incubating each enzyme with blood from two non-allergic healthy individuals and was generally observed down to 10 μg enzyme/ml. A significant histamine release at 1 μg enzyme/ml or less was therefore defined as a specific positive reaction.

In the second phase (STEP 2), all patients positive to one or more enzymes were investigated further. This was done by skin prick testing with the enzymes using the test batch (again) and, if available, the enzyme protein obtained by further purification of the test batch. Furthermore, as a model for everything in the test batch but the enzyme protein itself, fermentation broth from the wild-type *Aspergillus* and *Bacillus* strains grown using standard conditions and standard media were also tested (later referred to as the "wild-type model broth"). Please observe that the strains themselves were removed from the broth before using.

The histamine release testing was also repeated using the relevant skin prick testing material described above.

Subsequently double blind, placebo controlled food challenge (DBPCFC) was performed using commercially available enzyme products on separate days with the culprit enzyme(s) according to EAACI guidelines (Bindslev-Jensen et al., 2004).

4. Challenge material

For the food challenges were in all cases used the individually relevant commercial enzyme products as is. All the commercial enzyme products used have been tested and approved for food use in Denmark with specific usage limits. The enzyme products are generally widely used in different products in Denmark.

The maximum allowed dosages (in Denmark) of each enzyme product in each application have been used for the calculations of the dosages using a fixed intake of the

relevant food/beverage of 250 g in order to illustrate a worst-case situation. To further illustrate a worst-case situation it was in every case assumed that all enzyme activity was retained 100%, even though the enzyme product is largely inactivated and/or removed as a result of the food/beverage production process. The different relevant enzyme products were pooled into one active food challenge prepared especially for each patient, with the exception of the protease enzymes. Proteases may degrade other enzymes and therefore each protease was always given alone. This meant that in some cases a patient was given more than one active food challenge.

To blind the enzyme products they were placed in non-transparent cups with straws and dissolved in water (150 ml) and black currant juice (2 ml), the pH being close to neutral. The cups with content were frozen (−18 °C) immediately after the preparation and de-thawed just before use. Placebo challenges containing water and black currant juice only was used as well.

The protocol was approved by The Ethics Committee for Funen and Vejle Counties (Jr No: VF 20020198).

5. Results

Among the 400 tested allergic individuals, 387 (97%) did not react in the skin prick test when challenged with the selected enzymes using the test batches. Thirteen patients were positive, and their demographic data, sensitizations, skin prick testing and histamine release results obtained in STEP 1 are presented in Table 3. These patients reacted

Table 3
Demographic data, allergic diseases, allergies, skin prick testing and histamine release results of 13 patients having one or more positive skin prick test to an enzyme in STEP 1 and challenge results in STEP 2

Patient no.	Age	Sex	Allergic diseases	Allergies	SPT positive in STEP 1 (dilution)	HR positive in STEP 1 (μg/ml)	Challenge results in STEP 2
23	38	F	A, R, C, AE	Grass, mould, mite, milk	Enzyme 2 [(1:1)]		Negative to both active and placebo
26	37	F	A, R, C	Grass, mould, mite	Enzyme 1 [1:1]		Negative to both active and placebo
28	18	M	R, C, AE	Birch, grass, mugworth, dog, mite	Enzyme 2 [1:1] Enzyme 8 [1:1] Enzyme 11 [1:1]	 Enzyme 11 [1 μg/ml]	Two active challenges, Negative to both active and placebo
30	22	F	A, C, AE	Birch, grass, mugworth, dog, cat	Enzyme 6 [1:10]	Enzyme 6 [1 μg/ml]	Reaction to placebo, negative to active
63	57	F	A, R, C	Birch, grass, dog, mite	Enzyme 6 [1:1]		Negative to both active and placebo
66	42	F	R	Birch, grass, mugworth, cat	Enzyme 6 [1:1] Enzyme 7 [1:1] Enzyme 15 [1:1] Enzyme 19 [1:1]		Two active challenges, Negative to both active and placebo
83	45	F	R	Dog, cat, horse, mite	Enzyme 5 [±] Enzyme 13 [±] Enzyme 16 [±] Enzyme 17 [±]	 Enzyme 16 [0.1 μg/ml] Enzyme 17 [0.3 μg/ml]	Negative to both active and placebo
90	30	F	R, C, AE	Birch, grass, cat, poppy seed	Enzyme 6 [(1:1)]		Negative to both active and placebo
93	34	F	R, C	Grass, cod, shrimp	Enzyme 10 [(1:1)]		Negative to both active and placebo
134	46	M	R, C	Grass	Enzyme 10 [1:1] Enzyme 12 [(1:1)]	Enzyme 10 [1 μg/ml] Enzyme 12 [0.3 μg/ml]	Negative to both active and placebo
185	35	F	A, R	Birch, grass, mite, egg, milk	Enzyme 11 [1:1] Enzyme 12 [(1:1)]	Enzyme 11 [0.3 μg/ml] Enzyme 12 [0.1 μg/ml]	Negative to both active and placebo
367	27	F	A, R, C	Birch, grass, mugworth	Enzyme 1 [(1:1)] Enzyme 3 [(1:1)] Enzyme 6 [(1:1)] Enzyme 10 [1:10] Enzyme 11 [1:1] Enzyme 12 [1:1] Enzyme 13 [(1:1)] Enzyme 14 [1:1] Enzyme 15 [1:1] Enzyme 16 [1:1] Enzyme 17 [1:1]		Negative to both active and placebo
372	29	M	R, C	Birch, grass, dog, cat, mite	Enzyme 15 [(1:1)]		Negative to both active and placebo

F: Female, M: male, A: asthma, R: rhinitis, C: conjunctivitis, AE: a topic eczema. (1:1): The first SPT positive, but negative a repetition. ±: SPT negative (>3 mm), but flare seen in the area of the wheal.

Table 4
Positive skin prick testing and histamine release results in the 13 patients re-tested in STEP 2

Patient no.	Enzyme 6 test batch		Purified Enzyme 6		Enzyme 10 test batch		Purified Enzyme 10		Enzyme 12 test batch		Purified Enzyme 12		Enzyme 16 test batch		Enzyme 17 test batch		Purified Enzyme 17	
	SPT	HR	SPT	HR	SPT	HR	SPT	HR	SPT	HR	SPT	HR	SPT	HR	SPT	HR	SPT	HR
23																		
26																		
28																		
30	12	Neg	9	Neg														
63	7	Pos	5	Pos														
66	3	Neg	3	Neg														
83											0	Pos	0	Pos	3	Pos	0	Neg
90	0	Pos	0	Pos														
93					2	Pos	0	Neg										
134					2	Pos	0	Neg	3	Pos								
185																		
367	0	Neg	0	Neg	3	Pos	0	Neg	0	Neg	2	Pos	2	Pos	2,5	Pos	0	Neg
372			9	Neg														

SPT: skin prick test, all numbers indicate the wheel size in mm. HR: histamine release test. Pos: positive, Neg: negative. Only positive results are presented in the table (if the enzyme test batch elicited one or more positive SPT and/or HR, the results of the purified enzymes are also presented if tested, even if the results were negative in both SPT and HR).

with a positive SPT to various enzymes, most frequent to Enzyme no. 6, a glucoamylase (6 patients) followed by Enzyme nos. 10 and 11, two lipases (each 3 patients). In no cases, we found a positive SPT to dilutions below 1:10 of the stock solution. The patients with positive reactions in SPT and/or HR represented a variety of allergic diseases and all but one (pt 83) were sensitized to pollen(s).

The following enzyme test batches elicited no positive reactions in STEP 1: Enzyme nos. 4, 9, and 18.

In STEP 2, the 13 patients positive in STEP 1 were re-tested with the enzymes causing a positive reaction in the STEP 1 using the test batch(es) again, and, if available, the purified enzyme protein(s) and the wild-type model broth(s) as described in the test protocol.

In STEP 2, the following elicited no positive reactions in any of the patients: Enzyme no. 1 (test batch and purified), Enzyme no. 2 (test batch and purified), Enzyme no. 3 test batch, Enzyme no. 5 (test batch and purified), Enzyme no. 7 (test batch and purified), Enzyme no. 8 (test batch and purified), purified Enzyme no. 10, Enzyme no. 11 test batch, purified Enzyme no. 12, Enzyme no. 13 test batch, Enzyme no. 14 test batch, Enzyme no. 15 (test batch and purified), purified Enzyme no. 17 and Enzyme no. 19 (test batch and purified) (data not shown). Purified Enzyme no. 3, purified Enzyme no. 11, purified Enzyme no. 13, purified Enzyme no. 14 and purified Enzyme no. 16 were not tested as these materials were not available. As can be seen from Table 4, the remaining enzymes elicited positive reactions in SPT and/or HR in some cases; in only one case (Enzyme no. 6, a glucoamylase), however, a positive result was obtained with the purified enzyme proteins (obtained by chromatographic methods).

When testing the wild-type model broths most patients except no. 23, 26 and 63 were positive to the wild-type model broth obtained from *Aspergillus oryzae* and/or *Aspergillus niger* in either SPT and HR, but no patients were positive to wild-type model broth obtained from *Bacilllus* in neither SPT nor HR (data not shown).

The 13 patients were finally challenged in a double blind, placebo controlled protocol with either one or two (patients 28 and 66) different active challenges together with a placebo. One positive reaction (patient 30) to placebo was obtained. Thus, no positive challenges to the enzymes positive in SPT and/or HR in STEP 1 or STEP 2 were found (Table 3).

6. Discussion

Workers exposed to enzymes for use in the food industry may develop allergy to the enzymes via inhalation (Bernstein et al., 1994, 1999; Zober, 2002; Merget et al., 2001; Burstyn et al., 1998; Leser et al., 2001; Quirce et al., 2002; van Kampen et al., 2002), whereas sensitization to the enzymes by oral route in the consumers has not been described.

We investigated a possible clinical allergenicity of 19 enzymes used in the food industry. The enzymes were selected to reflect a wide range of enzymatic activities as well as enzymes being produced by both non-GMM (6 enzymes) and GMM (13 enzymes), see Table 2. The enzymes were tested both in vivo using skin prick testing and in vitrousing histamine release from human basophils. In only 13 of 400 allergic patients investigated (3%) a positive SPT was found, often accompanied by a positive histamine release. These 13 patients were further investigated using the test batch of the enzymes resulting in a positive initial testing again and including purified enzyme protein preparations, if available, and wild-type model broths. These preparations were used for SPT and HR. In 46 out of 55 reactions there was a concordance corresponding to 84% between SPT and HR. For the food challenges (DBPCFC) commercially available enzyme products were used.

The FAO/WHO report on assessment of allergenicity of foods derived from biotechnology (Aalberse et al., 2001) suggests using in vitro techniques such as measurement of specific IgE both when the product arises from an organism known to be allergenic in man (e.g. fish) (Hansen et al., 2004; Bindslev-Jensen et al., 2003) and when the product is produced in organisms not known to be allergenic. In the latter case, a targeted approach is recommended i.e. using sera from patients with a variety of allergies of different nature. We used skin prick testing with the test batches of the enzymes as the initial screening procedure since no commercial IgE methods are available and for practical reasons, since testing of 400 patients with 19 different subjects in histamine release would have been impossible. SPT in workers exposed to enzymes produced in *Bacillus* species has previously been demonstrated to be more sensitive than measurement of specific IgE (Bernstein et al., 1994).

The initial procedure was followed by retesting with crude and purified enzymes in the patients positive in the initial phase. Although some of the enzyme test batches also in the second phase elicited a positive response in SPT or HR, only one of the purified enzyme proteins (Enzyme no. 6, a glucoamylase) was positive in four of the patients (30, 63, 66 and 90). Patients no. 30, 66 and 90 were also positive in SPT or HR to the wild-type model broth from *A. niger* or *A. oryzae*. None of these patients were positive to moulds, where cross reactivity between Cladosporium or Alternaria has been described (Mari et al., 2003). The four patients were all positive to pollens and animal dander, but no data on possible cross reactions between *Aspergillus* and pollen resp. animal dander has been published – and presence of *Aspergillus* antigens in the purified Enzyme no. 6 was not investigated. Only one case report of allergic rhinitis to *Aspergillus* has been reported previously (Taj-Aldeen et al., 2003), whereas Allergic Bronchopulmonary Aspergillosis of course frequently is seen. In a large scale study involving more than 4000 patients, sensitization to *Aspergillus* measured by SPT was seen in 2.4% of the patients, a percentage increasing to 12.6% in patients sensitized to other fungi (Mari et al., 2003).

The reason for the positive findings in SPT and HR remains obscure; one suggestion would be that sensitization (route unknown) may occur in rare cases, but since none of the patients were positive to any of the commercial enzymes in the subsequent oral challenge using exaggerated dosages of the enzymes compared to normal daily intake, the findings are without any clinical relevance.

The allergen labelling regulations and the FAO/WHO decision tree all aim at protecting the allergic consumer – by using oral challenges with the enzymes as the final proof of non-reactivity these criteria has been fulfilled. These results were obtained using enzymes with a wide range of enzymatic activities in active forms, i.e. before they had been degraded by e.g. heat in the final commercial product thus adding a further safety factor to the findings.

There were no indications of cross-reactivity between the tested enzymes used in food and the main known allergens represented by the patients included in this study.

Considering the wide variety of enzyme classes and origins included in this study it is concluded that ingestion of food enzymes in general is not considered to be a concern with regard to food allergy.

References

Aalberse, S., et al. (2001). Evaluation of allergenicity of genetically modified foods. Report of Joint FAO/WHO Expert consultation on allergenicity of foods derived from biotechnology 22–24 January. Food and Agriculture Organization of the United Nation (FAO), Rome, Italy.

Bernstein, D.I. et al., 1994. Characterization of skin prick testing responses for detecting sensitiz to detergent enzymes at extreme dilutions: inability of the RAST to detect lightly sensitized individuals. Journal of Allergy and Clinical Immunology 94, 498–507.

Bernstein, J.A. et al., 1999. A cross-sectional survey of sensitization to *Aspergillus oryzae*-derived lactase in pharmaceutical workers. Journal of Allergy and Clinical Immunology 103, 1153–1157.

Bindslev-Jensen, C. et al., 2003. Assessment of the potential allergenicity of ice structuring protein type III HPLC 12 using the FAO/WHO 2001 decision tree for novel foods. Food and Chemical Toxicology 41 (1), 81–87.

Bindslev-Jensen, C. et al., 2004. Standardization of food challenges in patients with immediate reactions to foods – position paper from the European Academy of Allergology and Clinical Immunology. Allergy 59 (7), 690–697.

Burstyn, I. et al., 1998. Determinants of wheat antigen and fungal alpha-amylase exposure in bakeries. American Industrial Hygiene Association Journal 59 (5), 313–320.

Dreborg, S. et al., 1987. Results of biological standardization with standardized allergen preparations. Allergy 42 (2), 109–116.

FCC, 2001. Food Chemical Codex, Enzyme Preparations, General Requirements, Fifth ed., National Academy Press, Washington, pp. 146–152.

Hansen, T.K. et al., 2004. A randomized, double-blinded, placebo-controlled oral challenge study to evaluate the allergenicity of commercial, food-grade fish gelatin. Food and Chemical Toxicology 42 (12), 2037–2044, December.

JECFA, 2004. General Specifications and Considerations for Enzyme Preparations Used in Food Processing as proposed by the Joint FAO/WHO Expert Committee on Food Additives, Compendium of Food Additive Specifications, FAO Food and Nutrition paper No. 52 add. 9, Rome 2001, pp. 37–39.

Leser, C. et al., 2001. The "egg–egg" syndrome: occupational respiratory allergy to airborne egg proteins with consecutive ingestive egg allergy in the bakery and confectionery industry. Journal of Investigational Allergology and Clinical Immunology 11 (2), 89–93.

Mari, A. et al., 2003. Sensitization to fungi: epidemiology, comparative skin tests, and IgE reactivity of fungal extracts. Clinical and Experimental Allergy 33, 1429–1438.

Merget, R. et al., 1993. Diagnostic tests in enzyme allergy. Journal of Allergy and Clinical Immunology 92, 264–277.

Merget, R. et al., 2001. Baker's asthma due to xylanase and cellulose without sensitization to alpha-amylase and only weak sensitization to flour. International Archives Allergy and Immunology 124 (4), 502–505.

Quirce, S. et al., 2002. Soybean trypsin inhibitor is an occupational inhalant allergen. Journal of Allergy and Clinical Immunology 109 (1), 178.

Scheurer, S. et al., 2001. Cross-reactivity within the profillin panallergen family investigated by comparison of recombinzant profillins from

pear (Pyr c 4), cherry (Pru av 4) and celery (Api g 4) with birch pollen profiling Bet v 2. Journal of Chromatography B 756, 315–325.

Taj-Aldeen, S.J. et al., 2003. Allergic Aspergillus flavus rhinosinusitis: a case report from Qatar. European Archieves Otorhinolaryngology 260 (6), 331–335.

van Kampen, V. et al., 2002. Occupational airway sensitization due to subtilisin. Pneumologie 56 (3), 182–186.

Zober, A., 2002. Response to a case of occupational asthma due to the enzymes phytase and β-glucanase. Occupational and Environmental Medicine 59, 63–64.

nature biotechnology

Allergenicity assessment of genetically modified crops—what makes sense?

Richard E Goodman[1], Stefan Vieths[2], Hugh A Sampson[3], David Hill[4], Motohiro Ebisawa[5], Steve L Taylor[1] & Ronald van Ree[6]

GM crops have great potential to improve food quality, increase harvest yields and decrease dependency on certain chemical pesticides. Before entering the market their safety needs to be scrutinized. This includes a detailed analysis of allergenic risks, as the safety of allergic consumers has high priority. However, not all tests currently being applied to assessing allergenicity have a sound scientific basis. Recent events with transgenic crops reveal the fallacy of applying such tests to GM crops.

Genetically modified (GM) crops undergo rigorous safety assessment before being allowed to enter the market. One aspect of GM foods that has drawn a lot of public attention is the assessment of their potential allergenicity. Protecting people with food allergies against accidental exposure to allergens has become an important focus for food manufacturers and regulators responsible for all food safety. A significant focus of the food industry is to keep food products that are not intended to contain a major allergen (e.g., peanut, milk, eggs or wheat) from being contaminated with one of the major allergens. Likewise, the primary focus of the safety assessment for GM crops, as defined by the Codex Alimentarius Commission (**Box 1**)[1], is to prevent the transfer of a gene encoding a major allergenic protein (from any source), into a food crop that did not previously contain that protein.

The producers of GM crops and regulatory authorities focus on preventing avoidable increases in the risk of allergy in producing and accepting new GM crops. It should, however, be recognized that absolute avoidance of all risk is not achievable. Thus the assessment that has been developed focuses on avoiding risks that are predictable and likely to cause common allergic reactions.

Before discussing the details regarding the approaches used for assessing potential allergenicity of GM crops and the drawbacks of some steps, it is important to put the risks associated with food allergy into perspective. The prevalence of food allergy is not well established but is estimated to be around 6% in young children and 3% in adults[2].

Known potent allergenic foods like peanut or shrimp are not banned from the market, even though 1% of the population might develop allergic reactions upon exposure. In addition, market introductions in the recent past of novel foods like kiwi have resulted in the development of new allergies. Yet kiwi has not been removed from the market. Some of the major allergenic foods like fruits, nuts and fish are considered essential components of a healthy diet, and nobody would endeavor to deprive 99% of the population of these foods because 1% is at risk of developing food allergy. Instead, food labeling is used to help the allergic consumer avoid exposure to foods that cause their reactions. Similar arguments could be made for new crops developed either by conventional breeding or by genetic modification to, for example, help combat malnutrition in developing countries.

Furthermore, to date there is no documented proof that any approved, commercially grown GM crop has caused allergic reactions owing to a transgenically introduced allergenic protein, or that generation of a GM crop has caused a biologically significant increase in endogenous allergenicity of a crop[3]. However, the potential for the transfer of an allergen was illustrated in the 1996 case of transgenic soybeans into which the gene for a 2S albumin from the Brazil nut had been transferred to enhance the methionine content of animal feed. Although the protein had not previously been recognized as an allergen, a study sponsored by the developer of the crop, Pioneer Hi-Bred International (Johnston, IA, USA) during product development demonstrated IgE-binding with sera from Brazil nut–allergic subjects and positive skin prick tests to the transferred protein[4]. This protein is now known as the major allergen of the Brazil nut, Ber e 1. Despite being developed for animal feed only, the product was abandoned because of the obvious risk.

That experience provided guidance for development of the premarket allergenicity assessment process and demonstrated that specific, appropriate tests can prevent the transfer of a gene encoding a protein that might pose substantial risk. However, whereas absolute protection against all potential allergic reactions to a newly introduced protein can never be given, the allergenicity assessment of GM crops based on scientifically sound protocols should minimize the risks. It should be noted that some scientists and regulators have called for postmarket monitoring of GM crops to identify the development of new allergies associated with the crop. The full Codex guidelines[1], however, outlines the need for an effective premarket evaluation as the most effective tool to protect the public. There are technical, practical and economic issues that would need to be addressed in designing an effective postmarket monitoring system and are beyond the scope of this paper. Here, we focus on the scientific validity of protocols used in the premarket evaluation of the potential allergenicity of GM crops. In particular, we show how three tests that are commonly called for, and which have not been validated, can block development of potentially useful products.

[1]Department of Food Science & Technology, University of Nebraska, Lincoln, Nebraska, 68583-0955, USA. [2]Department of Allergology, Paul-Ehrlich-Institut, Langen, D-63225, Germany. [3]Division of Pediatric Allergy and Immunology, Mount Sinai Medical Center, New York, New York, 10029, USA. [4]Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, Victoria 3052, Australia. [5]Division Pediatric Allergy, National Sagamihara Hospital, Sagamihara, 228-8522, Japan. [6]Department of Experimental Immunology, Academic Medical Center, Amsterdam, 1105 AZ, The Netherlands. Correspondence should be addressed to R.E.G. (rgoodman2@unlnotes.unl.edu).

Published online 8 January 2008; doi:10.1038/nbt1343

Box 1 Risk assessment of genetically modified crops

The Codex Alimentarius Commission, under the FAO and the WHO, adopted guidelines in 2003 to harmonize the premarket risk assessment process for plants derived from biotechnology (GM plants) in the global market[1]. The guidelines were approved by the Codex Commission and are intended to guide countries in adopting consistent rules that provide a strong food safety evaluation process while avoiding trade barriers. Each new GM crop requires a premarket safety assessment to evaluate intended and unintended changes that might have adverse human health consequences caused by the transfer of the DNA (genes). The goal is to identify hazards, and if found, to require risk assessment and where appropriate develop a risk management strategy (e.g., do not approve, approve with labeling and/or monitoring, or approve without restriction).

The process is based on the science and requires the use of methods and criteria that are demonstrated to be predictive. New methods should be validated and demonstrated to enhance the safety assessment.

The framework to guide evaluation of potential safety issues requires detailed characteristics of:

- The GM plant and its use as food
- The source of the gene
- The inserted DNA and flanking DNA at the insertion site
- The expressed substances (e.g., proteins and any new metabolites that result from the new gene product)
- The potential toxicity and antinutritional properties of new proteins or metabolites
- The introduced protein compared with those known to cause celiac disease if the DNA is from wheat, barley, rye, oats or related grains
- The introduced protein for potential allergenicity
- Key endogenous nutrients and antinutrients including toxins and allergens for potential increases for specific host plants (DNA recipients)

Certain steps in the assessment require scientific assessment of existing information; others require experiments, in which case assay validation, sensitivity and auditable documentation are required.

Evolution of guidelines for allergenicity assessment of GM crops

Guidelines for allergenicity assessment of GM crops were published in three sequential documents that have been broadly recognized. The first comprehensive document was published in 1996 by the International Food Biotechnology Council (IFBC, Washington, DC) in collaboration with the International Life Sciences Institute (ILSI, Washington, DC)[5]. This was followed in 2001 by the UN Food and Agriculture Organization (FAO)/World Health Organization (WHO) consultation recommendations[6] and in 2003 by the Codex Alimentarius Commission guidelines[1]. The revised recommendations (FAO/WHO, 2001; Codex, 2003) were meant to correct shortcomings, although further clarifications are possible as we learn more about allergens and gain experience in test methods[7]. Several elements, however, are well established and have remained consistent throughout the three successive sets of recommendations.

All documents agree that introducing known allergens into a different species needs to be avoided as the primary risk is to those with existing allergies. If the source of the gene is a common allergenic food, or if the protein displays significant sequence identity with known allergens, the candidate protein should be evaluated for IgE binding using a sufficient number (e.g., for >95% confidence) of sera from patients allergic to the source of the allergenic food or to the sequence of the matched allergen. Those tests should reveal whether the gene codes for a yet unidentified allergen from a common allergenic source or whether IgE against known allergens cross-reacts with the homologous new transgenic protein. Another parameter included in all three guidelines is resistance of the candidate protein to digestion by pepsin, the rationale being that pepsin-resistant food proteins are more prone to induce systemic, severe symptoms. Perhaps more importantly, such stable proteins are also thought to be more potent sensitizers than proteins that are readily digested in the gut (that is, they are risk factors for induction of new allergies).

The IFBC-ILSI and FAO/WHO guidelines both used a decision tree to evaluate the risk of allergenicity[5,6], as reviewed previously[7]. The IFBC-ILSI document recommended *in vivo* clinical testing (skin-prick tests (SPT) and double-blind placebo-controlled food challenges (DBPCFC)), even if *in vitro* assays had demonstrated a lack of IgE binding for proteins from an allergenic source, if the protein sequence included as little as a single eight-amino-acid match to a known allergen. Even so, the FAO/WHO recommendations designated *in vivo* clinical testing as impractical and perhaps even unethical under most circumstances as a risk assessment tool, and suggested instead that negative serum testing alone, or in some circumstances SPT testing, but not food challenges, might be necessary to demonstrate a lack of risk. Another change recommended by the FAO/WHO[6] guideline was a six-amino-acid match to indicate a risk of cross-reactivity with allergens, rather than an eight-amino-acid match indicated by IFBC-ILSI[5]. Two additional new elements were added to the FAO/WHO (2001) recommendations: targeted serum screening—in which serum samples of patients allergic (or at least sensitized) to allergen sources broadly related to the source of the gene (sharing similar high taxonomic groups; e.g., monocots, dicots or arthropods) are used to detect or exclude potential cross-reactivity—and animal model testing. Targeted serum screening was recommended even when the transgenic protein did not demonstrate significant sequence identity to a known allergen or when the specific serum screening—using sera from subjects allergic to the source or the sequence-matched allergen—was negative. Animal testing was included despite recognition that validated models predicting risk of sensitization in humans do not (yet) exist.

The Codex Alimentarius Commission guidelines abandoned the risk assessment based on a decision tree and adopted a weight-of-evidence approach[1]. A decision tree was found to be too rigid in a situation where no single criterion is sufficiently predictive and evidence derived from several types of information, based on tests with different levels of validation, needs to be taken into account. Codex clearly emphasized the need to use scientifically validated testing, specifically removing the demand for nonvalidated animal tests and targeted serum screens and calling for validation of short-sequence matching routines. Instead, a 35% identity over an 80-amino-acid window was recommended as a sufficiently conservative prediction for potential cross-reactivity. These recommendations have not been accepted by some regulators. Clearly, the existence of multiple documents with diverging recommendations coming from different organizations has resulted in confusion and sometimes arbitrary inclusion of tests upon request from regulatory authorities. In some cases, regulators continued to base their judgment on nonvalidated (e.g., animal models) or even rejected (short-peptide matches) tests.

Assessment protocols

Here, we look at the scientific soundness of the principles and protocols for allergenicity assessment and present some recent case studies to

illustrate the inappropriateness of nonvalidated methods for allergenicity assessment, whether part of the FAO/WHO[5] recommendations or the Codex[6] guidelines. **Figure 1** outlines the Codex guidance's weight-of-evidence approach to evaluate the potential risk of food allergy.

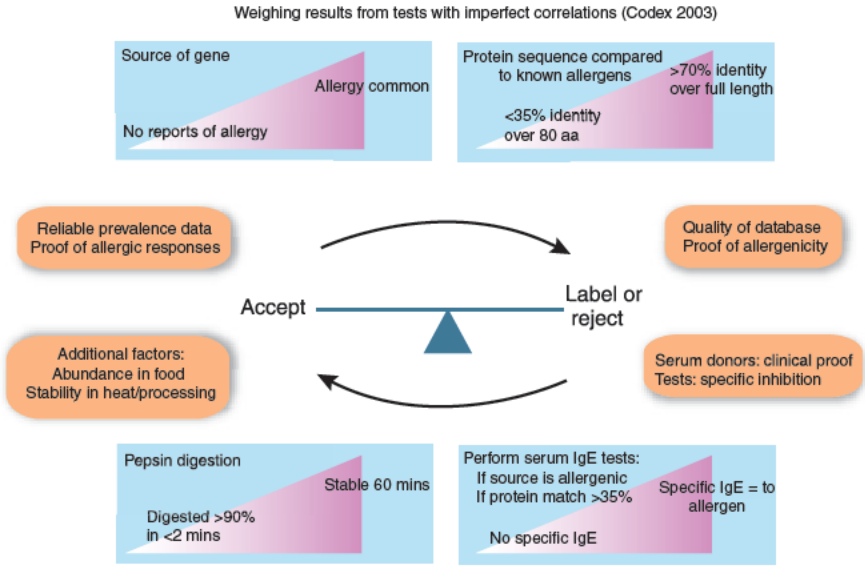


Figure 1 Schematic interpretation of the weight-of-evidence approach described by the Codex Alimentarius Commission Guidelines for Allergenicity Assessment in 2003 (ref. 1). In the figure, the four main areas of evidence are depicted with a graphic representation of the evidence representing maximum risk on the right (high side of the triangles). The weight of the evidence in each of the areas is influenced by the quality of the factors depicted in the yellow boxes. On the basis of the imperfect nature of the test methods available to distinguish between allergenic and nonallergenic proteins, scientific interpretation is necessary to reach a balanced and useful conclusion regarding the potential risks of allergy associated with each new food product.

Gene source. The process begins with an evaluation of the source of the gene. If the source of the gene encoding the new protein is a commonly allergenic food (e.g., peanut, hazelnut, hen's egg or cow's milk), a respiratory allergen (e.g., birch or grass pollen or house dust mite) or a contact allergen (latex), IgE-binding studies using sera from patients allergic to the source are required to ensure that the protein encoded by the gene does not bind IgE from those allergic to the source. For serum selection, demographic factors need to be taken into account. Both age and habitat have been shown to influence the molecular recognition profiles of specific IgE (**Box 2**). The number of sera needed is dependent on the degree of confidence considered necessary (largely a political and socioeconomic issue) and the prevalence of recognition of the hypothetical allergen. In other words, do we accept a 5% chance of an allergic reaction in 1% of the population allergic to the source or do we want to be more protective and only accept a 1% chance of a reaction in, for example, 0.01% of that population? Choosing to lower the risk requires a higher number of sera.

If the source of the gene rarely causes allergies, it would be difficult or impossible to find enough qualified serum donors to perform statistically valid tests. However, that also means the number of individuals in the population who would be at immediate risk of reactions if the protein were an allergen would be small. In such cases, the number of individuals is not as important as the specificity of the test and evidence of clinical relevance of the allergenic source.

Bioinformatics. The amino acid sequence of all transferred proteins, regardless of the source, are to be compared with known allergens by FASTA or BLAST algorithms to determine if any identity match is sufficiently high to suspect that the protein might cause allergic cross-reactions. This is not meant to be a stand-alone test, but rather to identify proteins that would require serum testing, using donors with specific allergies to the source of the sequence-matched allergen to evaluate potential IgE binding. If the identity match is high (e.g., >70% over most of the length of the protein), the potential for cross-reactivity is

Box 2 Spaniards are different from Dutchmen

Exposure to allergen is an essential prerequisite for sensitization[47]. An exception to this rule is cross-reactivity: for example, exposure to birch pollen can induce allergy to apple, cherry and hazelnut[27,48,49]. This is typically seen in those areas of the world where birch pollen exposure is high, such as The Netherlands. In the absence of birch pollen, apple allergy also exists, for example, in Spain. In a recent European multicenter study, almost 400 people allergic to apple from four countries were compared to identify potential cross-reactive causes[48]. As expected Dutch, and others (Austrian and Northern Italian) individuals were allergic to apple because they were allergic to birch pollen. IgE binding the major birch pollen allergen Bet v 1 cross-reacted with the homologous major apple allergen Mal d 1. Symptoms induced by Mal d 1 were almost exclusively mild and restricted to the oral mucosa. Spanish participants had not been exposed to birch pollen and were shown to be sensitized to a non-pollen–related allergen identified as a lipid transfer protein (Mal d 3). Although the majority exclusively have mild symptoms in the oral cavity, it was demonstrated that IgE against lipid transfer protein is a significant risk factor for the development of severe systemic symptoms, as were observed in ~25% of the Spanish individuals[48]. This study clearly illustrates that the outcome of allergenicity assessment of GM crops using serum samples of patients with largely identical clinical symptoms upon consumption of apple is strongly influenced by the geographic origin of the patients. Spanish are simply different from Dutch apple-allergic patients due to differences in exposure or other local environmental factors; variations in genetics of these populations cannot account for the marked differences. In the former case, assessment will focus on non-pollen–related apple allergens; in the latter, on birch pollen–related allergens. Similar patterns have been reported for cherry allergy with Pru av 1, a homolog of Bet v 1, being the dominant allergen, compared with lipid transfer protein in cherry, peach and hazelnuts as the primary allergen in the Mediterranean areas[27,49]. These studies highlight the need for good patient characterization and selection before the use of their sera in allergenicity assessment protocols, as differences in the prevalence of IgE sensitivity is possible in the same foods, in different populations.

high and the risk would probably be close to that posed by the matched allergen. Matches sharing between 50% and 70% overall pose a moderate risk of cross-reactivity and should be tested for IgE binding. If the match is <50% identical, the risk of cross-reactivity is expected to be low[8]. Even so, a conservative threshold value of 35% identity over any 80-amino-acid segment of the transferred protein contained in both the FAO/WHO[6] and Codex documents[1] was intended to identify conserved gene segments representing functional motifs, which might retain conformational epitope structure as well. Proteins with higher matching identities (e.g., >35% identity) are recommended for testing of IgE binding.

On the basis of literature searches, only a few examples of endogenous proteins from sources suspected of cross-reactivity demonstrate significant IgE cross-reactivity for proteins sharing between 35% and 50% identity over the entire length of both proteins, and quantitative IgE binding and basophil histamine release (an *ex vivo* test of the circulating effector cells triggered to release histamine by IgE cross-linking) demonstrate only partial reactivity[7]. The lack of known examples of cross-reactivity associated with proteins sharing only 35% identity over 80 amino acids suggests the criterion is too conservative as it would overpredict potential cross-reactivity. One alternative is to focus on overall sequence alignments, as suggested by Ladics *et al.*[9]. Another alternative would be to increase the percent identity for the 80-amino-acid window closer to a level (possibly >50% identity) where there are examples of at least weak *in vitro* cross-reactivity using sera from individuals having allergic symptoms to the sources of both proteins[10].

The bioinformatics step is relatively straightforward and should markedly reduce the risk of transferring even a minimally cross-reactive protein. However, some allergens that may be matched are rarely noted as causing allergies and it would be virtually impossible to identify appropriate serum donors for a well-powered study. In such cases, the risk of potential allergy to the population from that protein is likely to be extremely low and regulators may be willing to waive the requirement for IgE testing. Choosing the appropriate allergen database to search is vital for a reliable sequence comparison[10]. AllergenOnline (http://www.allergenonline.com) is the only database that is currently fully peer-reviewed regarding evaluation of published evidence of allergenicity. Other databases are available and the alternative of searching the National Center for Biotechnology Information (NCBI; Bethesda, MD, USA) nonredundant database, with key-word limits can provide more updated sequences, but lacks an accurate screening method for relevance[10].

Although there are not yet any publications reporting validation of the approach using 35% homology over an 80-amino-acid window (or >50% overall homology) to predict likely cross-reactivity, it is clear that it is an improvement over methods using sequence homology over 6 or 8 amino acids[11]. These short peptide matches have not been validated as predictive tools and should be rejected on the basis of extremely high numbers of false-positive hits[11–14]. The eight-amino-acid match was originally selected without evidence of predictability based on the idea that it would represent both a theoretical B-cell epitope as well as a minimum size for a conserved T-cell epitope[5]. Stadler and Stadler[13] reported that a 6-mer match resulted in more than two-thirds of all proteins in Swiss-Prot being predicted to be allergens, and >40% of the human genome being predicted as such. Obviously, the use of short amino matching searches (6–8 mer) is not a useful approach for allergenicity assessment, but it has never been truly renounced. Consequently, a few regulatory authorities sometimes still require bioinformatics analyses based on 6-mer matches (**Box 3**).

Serum IgE binding. Serum IgE testing to evaluate proteins from an allergenic source, or proteins with sequence identity (e.g., >35% over an 80-amino-acid widow or >50% overall) to a known allergen works very well if performed properly[15]. Appropriate positive and negative control proteins or extracts of the allergenic source material are required to demonstrate assay validity. The positive test sera must be from clearly diagnosed allergic subjects who react to the gene source or sequence-matched allergen and its source. Negative control sera would typically include donors with allergies to other unrelated proteins as well as non-allergic subjects. A few relatively well-controlled studies have been used to evaluate GM crop safety[4,15,16], although the relevance of donor selection has not always been clear[15].

The design and interpretation of assays for specific IgE can be complex. Potential confounding factors include the molecular appearance of the protein (e.g., monomeric versus multimeric, proper folding or misfolding, presence or absence of disulfide bonds, presence or absence of N-linked glycans) and abundance of the protein in the source material (that is, sensitivity). The test material must be representative of the form available in the GM food source. The tests should be capable of detecting IgE binding to linear and conformational epitopes, sometimes requiring two separate assays (e.g., under reducing and native conditions). Demonstration of specificity of binding requires replicate samples with sera exposed to appropriate inhibitors.

Moreover, it must be recognized that there are no absolute thresholds of serum IgE binding that provide absolute measurement of safety or risk. Positive IgE tests without clinical relevance are common in clinical

Box 3 Short peptide match: a lot of work for nothing

Pioneer Hi-Bred International and Dow AgroSciences (Indianapolis, IN, USA) developed a GM maize product containing the gene encoding Cry1F, from *Bacillus thuringiensis*. The product was approved for sale in the United States and Canada following full regulatory studies, including assessment of the potential allergenicity of the protein based on Codex guidelines[1]. The protein produced from this gene is toxic to lepidopteran larval pests, such as the European corn borer, but not to mammals[16]. The gene is from an organism not known to cause allergies. The sequence is not significantly identical to any known allergen based on overall FASTA alignment. It is <35% identical to any 80-amino-acid segment of known allergens, which is the primary alignment criterion recommended by Codex[1]. Because of regulatory requests from Taiwan, an additional bioinformatics comparison was performed to identify any six-amino-acid matches with allergens. There was a single six-amino-acid match to the house dust mite allergen Der p 7 (ref. 16). The protein does not have any other alignment similarity to Der p 7, yet regulators from Taiwan required human allergic serum IgE testing to evaluate potential cross-reactivity. The results of the serum IgE test demonstrated a lack of IgE binding to Cry1F using sera from allergic subjects who had clear IgE binding to Der p 7 (ref. 16). The results satisfied the regulators and the product was approved. However, the tests were expensive and there is always a chance of obtaining a weak-positive IgE binding result. Even the slightest amount of binding would likely have led to extensive *in vivo* testing, but would have been unlikely to demonstrate a risk of an allergic response in consumers as at least two IgE binding sites and high affinity are required to effectively cross-link mast cells and trigger an allergic response (as discussed in reference 11).

practice (e.g., due to the presence of cross-reactive IgE to plant N-glycans). To avoid potentially confounding test results, developers may want to remove glycosylation sites before introducing the new gene unless the glycan is needed for functionality. Serum from individuals with strong carbohydrate-specific IgE antibodies should be avoided for GM assessment to ensure selection of appropriate donors who should have IgE directed against peptide epitopes rather than carbohydrate. Otherwise, carbohydrate-binding sera would lead to designating most glycoproteins as an allergenic risk, although it is widely accepted that the glycans are unlikely to cause clinical food allergy[17,18]. In the event the transgenic protein is glycosylated, alternative testing may be required to evaluate glycan structure or if IgE binding is demonstrated, the relevance should be tested by basophil histamine release or *in vivo* allergen testing. Diagnosing allergic disease requires a holistic evaluation of diet, symptoms, SPT and/or specific IgE and elimination diet or challenge test[19]. Likewise, interpretation of IgE binding to GM proteins requires judgment. Strong, specific binding to the protein using appropriate donors should be taken as evidence of risk. However, low levels of binding that are not clearly specific and close in affinity to the suspected cross-reactive allergen may not indicate significant risk. If results are equivocal, SPT or challenges might be necessary to demonstrate the relevance of low levels of apparent specific IgE binding.

Stability in pepsin and abundance. The ability of the new protein to withstand digestion by pepsin is evaluated as a potential risk factor of allergenicity[20,21]. Several potent food allergens are known to be very stable in an *in vitro* pepsin digestion assay, whereas it is thought that most dietary proteins are readily digestible[22]. However, some proteins not known to cause significant food allergies are also stable[23]. And some proteins known to cause food allergy, especially those inducing only oral allergy syndrome—mild tingling or itch in the mouth, without substantial edema—are relatively labile[24]. Thus far, food allergens from this last category are mainly found among cross-reactive allergens, where primary sensitization occurs by inhalation (e.g., pollen or latex). These are therefore usually not designated to be 'true' food allergens[25]. Such proteins are likely to pose little risk to consumers if expressed at low abundance in crops.

Some very stable proteins such as thaumatin-like proteins from apple and grape rarely cause allergy or possibly only mild reactions[26], whereas others, like the lipid transfer proteins from a variety of sources, are very stable and may frequently cause severe reactions[26,27]. Some of these stable proteins are inducible pathogenesis-related proteins and expression is variable in foods, which may complicate their recognition as allergens[28]. There is also evidence that some important pepsin-labile allergens become more stable with minor shifts in pH (e.g., from pH 2.5 to 2.75 for codfish parvalbumin)[29]. Although the increased stability at moderate stomach pH values may help explain the allergenicity of some of these proteins, the use of standard pepsin stability testing at pH 1.2 or 2.0 still has a good demonstrated predictive value[30].

An additional risk factor for food allergy is the abundance of the protein in food, as many major food allergens account for >1% of the protein in high-protein allergenic foods[20]. Others, such as lipid transfer proteins and parvalbumins are less abundant. Abundant, pepsin-stable proteins are more likely to survive digestion in sufficient quantities to facilitate sensitization and become significant food allergens. The typical quantity consumed of specific foods would be expected to have an impact as well, so nonabundant, stable proteins may be potent allergens if a large amount of food is consumed. However, additional scientific data would be required to establish completely objective criteria for acceptance or concern based on stability and abundance. Currently the results are judged relative to common, potent food allergens.

Box 4 Mission impossible: evaluation of changes in endogenous 'hypo-allergenicity'

A transgenic herbicide-tolerant rice, Liberty Link-rice (LLRICE62), was produced by Aventis CropScience (now Bayer CropScience, LP, Research Triangle Park, NC, USA), by inserting the gene for phosphinothricin-*N*-acetyltransferase (PAT) from a bacteria that has not been reported to be allergenic, nor does it share significant sequence identity with any known allergens. The nonglycosylated PAT protein is rapidly digested by pepsin under standard conditions[50]. On the basis of these characteristics, there is no need to test IgE binding to evaluate the potential allergenicity of the PAT protein. US regulators approved the product in 1999 (http//www.agbios.com/dbase.php). However, because rice has been reported (rarely) to cause allergic reactions in humans, the developer performed an *in vitro* IgE binding study of LLRICE62 to compare endogenous allergenicity to a nontransgenic cultivar after their interpretation of the IFBC-ILSI recommendations and based on historical questions from regulatory agencies (pre-1999). Because true (challenge-positive) rice-allergic individuals cannot easily be found, sera of food-allergic subjects with rice-specific serum IgE or skin test–positive reactions to rice extract, or individuals with clinical histories suggestive of rice allergy were used. However, rice-food allergy was not confirmed by food challenge. These individuals were probably sensitized to grass pollen or inhaled rice flour and may be unaffected when ingesting rice based on a paucity of published cases of proven rice allergy and our experiences[51,52].

The unpublished study (personal communication, Donna Mitten, Bayer CropScience, data reviewed by R.E.G.) revealed no significant differences in IgE binding and allergen content between the GM and a genetically similar traditional rice variety. The value of a study based on sera of patients with unconfirmed rice allergy is questionable. Regardless, Canadian authorities approved LLRICE62 in 2006 having been satisfied with the assessment of potential allergenicity that included an evaluation consistent with current guidelines in addition to the results of the serum study (http://www.hc-sc.gc.ca/fn-an/gmf-agm/appro/nf-an90decdoc_e.html). It can be argued that it is not justified to evaluate GM crops for potential changes in endogenous allergenicity for a food with extremely low allergenicity because results will generally be meaningless.

However, the story does not end here. Recently, a small amount of seed from a similar GM event (LLRICE601) was discovered in a commercial 'nontransgenic' rice variety. In order to quickly respond to regulatory requests for information about studies of LLRICE601 safety, Bayer CropScience considered the earlier Canadian request for LLRICE62 and decided to perform a similar study if feasible. However, the original sera used for the previous study were no longer available (personal communication, Donna Mitten). Communications with clinical allergists (including M.E., D.H., H.A.S.) in Australia, Japan, Korea, Taiwan and the United States (by R.E.G) indicated the extreme difficulty in obtaining serum donors with clinically defined allergy to rice (as food), and the study was not performed. But, because people who are allergic to rice should avoid eating it, and so few are allergic, it is not clear that there would be any value in performing such a study.

GM crop allergenicity assessment—what is not (yet) useful?

As the assessment of the allergenicity of GM crops has evolved, scientific evaluation of some tests and criteria included in earlier guidance has demonstrated that some specific approaches are not (yet) particularly useful (e.g., six- to eight-amino-acid matches, targeted serum tests). Additionally, some new approaches have been espoused that are not sufficiently validated in terms of predicting allergenicity for use in regulatory decision making, although in some cases continued research may be warranted.

Endogenous allergenicity. If a transgene is transferred to a commonly allergenic food, it appears logical to monitor the influence on endogenous allergenicity, which was recommended by the various guidelines on GM crop allergenicity assessment[1,5,6]. These documents, however, have not addressed the level of change that would be (un)acceptable. Several studies have been carried out comparing endogenous allergenicity of nontransgenic and GM varieties. Monsanto (St. Louis, MO, USA) performed an evaluation of herbicide-tolerant soybeans using sera from soy-allergic subjects[31], and also of a potential herbicide-tolerant GM wheat product using sera from ten subjects allergic to wheat. In a comparison of IgE binding, neither study demonstrated any significant differences between the GM crop and non-GM controls (R.E.G., poster presentation, World Allergy Organization meeting, Vancouver, BC, 2003). A study by Lehrer and Reese[32], commissioned by Pioneer Hi-bred International, compared conventional and GM high–oleic acid soybeans using sera from five individuals selected for high IgE binding to soybean extract. A radioallergosorbent-inhibition (RAST-inhibition) assay demonstrated similar IgE binding results between the GM and non-GM varieties. However, what is the risk and what should be done if statistically significant differences are detected?

Serum IgE binding and histamine release were tested in a comparison of ten varieties of Roundup Ready soybean (GM) developed by Monsanto and eight cultivars of non-GM soybean[15]. IgE-inhibition tests demonstrated up to fourfold differences in IgE-binding potencies across both the GM and non-GM varieties, but overall the GM and non-GM varieties were not significantly different. That study illustrated that a head-to-head comparison of a pair of randomly selected soybean varieties may lead to statistically significant differences, even though the apparent allergenicity of the individual varieties falls within the range of responses to several commercially available non-GM soybeans. Apart from the fact that serum samples used in this study originated from subjects that were negative to soy by food challenge (or were not challenged), the variable IgE binding results clearly highlight an aspect that should be taken into account when evaluating effects on endogenous allergenicity: natural variation of allergenicity of available food crops due to differences in the genetics of commercial varieties, and interactions with the environment (e.g., nutrient availability, differences in moisture, temperature, plant pathogens). It is unreasonable to be more stringent toward GM crops with respect to changes in endogenous allergenicity than can already be accounted for by natural variability. Basing judgment on statistical significance alone has no clinical meaning if natural variability is larger. Importantly, the whole discussion about endogenous allergenicity has limited relevance because patients allergic to the food will (should) avoid eating it anyway, GM or not, to avoid allergic reactions.

The soy study results[15] suggest that there is wide variation in IgE binding to different varieties of the same species of non-GM crops, but few studies have been performed to study the question in a systematic way. Various groups have addressed differences in allergenicity between non-GM apple cultivars, focusing on two major apple allergens, the birch pollen–related allergen Mal d 1 and a lipid transfer protein, Mal d 3. Differences in allergenicity have been found by IgE-binding and IgE-inhibition studies, immunoassays for quantifying allergens, *in vitro* basophil histamine release and genomic sequence variability, but also by SPT and DBPCFC as illustrated below.

Sequence variability, possibly translating into differences in allergenicity, has been recently reported for both Mal d 1 and Mal d 3 in different apple cultivars[33]. Most studies focusing on differences in allergenicity of apple cultivars have used IgE-based binding (*in vitro* and *in vivo*) as an endpoint. A recent study has evaluated IgE binding and SPT reactivity as well as measuring Mal d 3 content, comparing ten cultivars of apples[34]. The Mal d 3 content varied more than

Box 5 A controversial nonvalidated animal model

A gene encoding an α-amylase inhibitor 1 (αAI) was transferred from kidney bean to field peas to make peas resistant to a bruchid storage beetle[53]. Because of the recommendation for animal model tests by the FAO/WHO[6], the developer tested the product in a mouse model using repetitive intragastric sensitization followed by intratracheal challenge[54]. This model had not previously been used to predict allergenicity of food proteins and we are aware of no other studies that have used an airway challenge or measure of pulmonary cellular infiltration to evaluate food allergenicity. The test results demonstrated stronger eosinophil accumulation in the lungs in mice sensitized and challenged with the GM pea (or αAI from the pea), compared with the kidney bean[54]. This supports increased Th2 inflammation, but not necessarily IgE-mediated allergy. The report described structural differences of the *N*-linked glycan on αAI expressed in peas compared to kidney bean. There was also evidence of different proteolytic processing of the C terminus of the protein. The authors concluded that differences in post-translational proteolytic processing were responsible for the apparent enhanced immunogenicity of the GM product[53].

The mechanism leading to the altered response in mice is not clear, but more importantly, the model has not been widely tested with allergenic and nonallergenic proteins as would seem necessary based on Codex guidelines[1]. In the case of the GM αAI pea, the differences found in glycan structure and protein processing would have been more appropriately investigated by human serum testing to evaluate IgE binding using serum donors with allergies to legumes if regulators wished to have testing beyond the bioinformatics, pepsin digestion and characterization of the protein.

Despite the fact that no scientific evidence was provided for an increased risk of IgE-mediated food allergy in humans, the study aroused a storm of negative publicity for GM crops, being an allergy risk[55]. Although the developer did not report results of a bioinformatics evaluation of the protein, in our hands a FASTA search of AllergonOnline (http://www.AallergenoOnline.com/), version 7.0, revealed one match of 41% identity over an 80-amino-acid segment to peanut agglutinin precursor, a putative allergen. The overall identity was 34.5%. Although this low level match is not likely to indicate cross-reactivity, it is above the Codex criterion. The data suggesting that peanut agglutinin is an allergen should be evaluated in making a final decision on whether to perform human serum-IgE testing, before any regulatory decision to approve the GM crop. In any event, data from a mouse model should not be relied upon to predict allergenicity.

sixfold on a dry material basis across cultivars. The mean wheal area resulting from SPT of the highest Mal d 3 content apple variety (~55 mg/g) was significantly higher (~threefold) than the mean wheal area for the two varieties with lower concentration (~10 mg/g) of Mal d 3. One may infer from the SPT results and Mal d 3 quantities that the cultivar with the highest levels of the allergen (Starking) is probably three times more allergenic on a gram basis than those with lower levels (e.g., Golden Delicious). Similar studies[35] were carried out with 88 apple cultivars focusing on both Mal d 1 and Mal d 3, although not all the results are published (R.v.R., unpublished data). In both cases, differences in allergen content differed up to 100-fold between the extremes, both in allergen quantification and IgE-inhibition assays. Some of these differences had been observed in SPT and DBPCFC testing, with about tenfold differences between individual cultivars. These detailed studies demonstrate the wide range of natural variability of allergenicity in a common non-GM food.

Similar tests of soybean varieties by *in vivo* skin reactivity and *in vitro* IgE binding of ten soy cultivars found up to sixfold differences in IgE-binding potencies[36]. Apart from differences between cultivars, natural variability in allergenicity can also occur due to harvest timing and storage conditions[37,38]. Even between individual apples from a single cultivar and harvest, up to tenfold differences in allergenicity have been reported[39]. Yet clinicians and food safety experts do not recommend avoiding certain apple or soybean varieties, nor is there evidence of significant differences in clinical reactivity for the allergic consumer.

Overall, these studies demonstrate the need to establish natural variability of allergenicity of non-GM crops before demanding evaluation of changes in endogenous allergenicity of GM crops. Nevertheless, some regulatory authorities have interpreted the guidelines so broadly that they demand evaluation of changes in endogenous allergenicity of foods for which it is virtually impossible to find sufficient truly allergic patients for a well-powered study (**Box 4**).

Of course, in cases where there are specific reasons to suspect a major impact on expression levels of endogenous allergens, special attention has to be given to evaluating allergenicity. This can, for example, be the case when a transcriptional activator is inserted or the transgene is inserted in the coding region for an allergen. Such events should however, not go unnoticed by the developer of a GM crop as detailed molecular characterization of the insert and the protein as well as protein function are required by Codex[1] (**Box 1**).

Targeted serum screens. The FAO/WHO[6] recommendation for broadly targeted serum screens specifically stated that if the source of the transferred gene was a monocotyledonous plant (class Liliopsida), serum should be taken from 50 individuals with allergies to diverse monocot sources (e.g., some allergic to grass pollen, maize, rice or dates) to identify potentially cross-reactive allergens. However, in the Codex guidelines, this was recognized as unlikely to be predictive[1]. There are four or five structural protein families (prolamins, Bet v 1-relatives, cupins and profilins) with representative clinically cross-reactive allergens from taxonomically diverse sources[40]. Although a few individuals react to material from sources as diverse as representatives of an order (e.g., Fabales) or even higher group, most clinically important cross-reactions are elicited by material from within the taxonomic family (e.g., Fabaceae), tribe (e.g., Phaseoleae) or, more commonly, genus (e.g., *Phaseolus*)[41]. Differentiating between clinical cross-reactivity, cosensitization and irrelevant IgE binding (low affinity or binding to cross-reactive carbohydrate determinants) is often complicated as clinical reactivity is rarely measured, rather some level of skin prick sensitivity or direct *in vitro* IgE binding is used to define cross-reactivity and this is likely to overestimate clinical reactivity[10,42,43]. Although validated specific serum tests with samples from clinically well-characterized subjects allergic to the source of the gene—or allergic to a sequence-matched allergen—should be useful when the need is indicated, targeted testing is unlikely to provide reliable data for the assessment.

Animal models. The FAO/WHO[6] recommendations called for evaluating each new GM crop with studies in two separate species of animals and/or using two routes of sensitization in one species, even though the panel recognized that no current animal model is predictive of allergenicity in humans. There are still no validated animal models for predicting allergenicity to food proteins, even though many models have been successfully applied to dissect mechanisms of allergic responses and potential changes due to modification of the allergenic proteins[44,45]. Even though many authors recognize that different animal models respond to specific proteins differently[46], they still suggest using animal models in the safety evaluation process for GM crops. On the basis of the paucity of correlative data between any one animal model and human food allergenicity, and the complex genetic diversity that predisposes subjects to allergy, it is not clear that any animal model could be useful in predicting the potential allergenicity in humans of a novel protein or GM crop. It is also not clear how one might combine results from two animal model tests to produce a predictive result. An unpublished study coordinated by the ILSI-Health and Environmental Sciences Institute (Washington, DC) reported results from a multi-laboratory test of the most commonly used mouse strains (BALB/c, C3H/HeJ, A/J and BDF-1) using commonly recommended protocols to evaluate IgE and allergic responses to identical samples of common potent allergens of peanut (Ara h 1 and Ara h 2) and milk (beta-lactoglobulin) compared with relatively nonallergenic proteins of spinach (RUBISCO) and soybean (lipoxygenase). The responses to the potent allergens were equivalent or weaker than responses to the weakly or nonallergenic proteins (Thomas, K. *et al.*, 2005 annual meeting poster, American Academy of Allergy Asthma and Clinical Immunology).

Box 6 Balb/c mice no substitute for human IgE recognition evaluation

A gene encoding amarantin was transferred from *Amaranthus hypochondriacus* into maize[56]. Although the protein was digested in the pepsin assay, comparing the sequence to known allergens identified a number of 6-, 7- and 8-amino-acid matches to known allergens[56]. Although noting overall homology to some allergenic proteins, the developer decided to use animal models to evaluate the allergenicity of the GM maize[56]. Comparing the amarantin sequence by FASTA demonstrated up to 70% identity over an 80-amino-acid segment to known allergens and >40% identity for overall alignments to a number of important 11S globulin allergens[7]. Clearly, this should have set off an alarm calling for serum IgE testing, if not immediately convincing the developer that the protein was too risky to transfer. Instead, the immunogenicity of the product was tested in BALB/c mice, with results demonstrating no significant response and the authors suggested there was no significant risk of allergy[56]. Although it is not clear if this potential product has been submitted for regulatory review anywhere, the Codex guidelines (2003) indicate that the amarantin-containing maize would require serum IgE testing with sera from at least a number of buckwheat- and/or Brazil nut–allergic subjects and possibly others.

On the basis of current knowledge, therefore, we recommend continuing research to evaluate potentially predictive animal models but caution against testing potential products at this time as there is no scientific validation demonstrating predictive values that are acceptable for risk evaluation. This opinion is also reflected in the recommendations of the Codex Alimentarius Commission[1]. Producers as well as regulators are sometimes confused about which recommendations to follow. This is illustrated by two cases in which developers of GM crops used animal models to evaluate potential allergenicity. In one, case results were interpreted as demonstrating likely allergenicity of the GM crop (**Box 5**) and in another case to suggest absence of allergenicity, even though there is a strong potential for cross-reactivity based on bioinformatics (**Box 6**). In either case, there is no scientific justification for these conclusions.

Conclusions

The current safety assessment outlined in the Codex guidelines (2003)[1] is based on the current state of knowledge regarding food allergens and risk, and is therefore well-suited to evaluate the potential for increased risk in allergenicity of GM crops compared with the risk of allergy from the conventionally bred crop varieties. The weight-of-evidence approach was adopted in part as it was recognized that there are exceptions to each component in the process. Thus, each product must be reviewed on a case-by-case basis and experienced scientists must be able to interpret results in aggregate. Key elements of this weight-of-evidence assessment are illustrated in **Figure 1**:

- Source of the gene: common allergen or not?
- Bioinformatics: sequence searches for matches of >35% identity over 80 amino acids (or of >50% overall identity for more realistic risks).
- IgE-testing: does the introduced protein bind IgE-antibodies?
- Stability testing: is the expressed protein highly resistant to digestion by pepsin?
- Abundance: is the protein abundant in the food (and stable)?

The premarket assessment recommended by Codex provides a mechanism to intercept GM crops that are likely to increase the risk of food allergy, as demonstrated by the identification of the Brazil nut 2S albumin transferred to maize, and the amarantin transferred to maize (**Box 6**) as proteins that would likely present significant health risks for specific populations of allergic consumers. The premarket screening process helps to avoid possible severe reactions in unsuspecting allergic consumers and also prevents subsequent costly food and seed recalls that would be needed to prevent additional reactions.

There is no scientific justification for inclusion of the following tests in allergenicity assessment because their predictive values have not been validated:

- Bioinformatics: short-peptide matches resulting in random false-positive hits.
- Animal models: useful for mechanistic studies but not applicable for prediction of human sensitization to food.
- Endogenous allergenicity: natural variability needs to be taken into account first.
- Targeted serum screens: potentially high rate of false-positive and low probability of true-positive results.

Demanding inclusion of such nonvalidated tests can lead to the rejection of safe and beneficial products, excessive costs and, potentially, disruption of trade without any further reduction of risk. Importantly, the use of inappropriate tests such as unvalidated animal models in place of more appropriate tests could lead to the introduction of a product that does pose substantial risk for a group of allergic consumers.

ACKNOWLEDGMENTS

The preparation of this article was conducted with a contribution of the University of Nebraska Agricultural Research Division, supported in part by funds provided through the US Department of Agriculture. Additional support was provided by the Food Allergy Research and Resource Program. Mention of a trade name, proprietary products or company name is for presentation clarity and does not imply endorsement by the authors. R.E.G. acknowledges Bayer CropScience for providing funds to support research for evaluating methods to compare endogenous allergenicity of crop varieties through research at the University of Nebraska. S.V. acknowledges Monsanto Company for supporting studies on the bio-variability of the allergenic potential of soybean varieties in comparison to transgenic lines.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Codex Alimentarius Commission. Alinorm 03/34: Joint FAO/WHO Food Standard Programme, Codex Alimentarius Commission, Twenty-Fifth Session, Rome, 30 June–5 July, 2003. Appendix III, Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants and Appendix IV, Annex on the assessment of possible allergenicity, pp. 47–60 (2003).
2. Sampson, H.A. Food allergy—accurately identifying clinical reactivity. *Allergy* **60** S79, 19–24 (2005).
3. Taylor, S.L. Review of the development of methodology for evaluating the human allergenic potential of novel proteins. *Mol. Nutr. Food Res.* **50**, 604–609 (2006).
4. Nordlee, J.A., Taylor, S.L., Townsend, J.A., Thomas, J.A. & Bush, R.K. Identification of a Brazil-nut allergen in transgenic soybeans. *N. Engl. J. Med.* **334**, 688–692 (1996).
5. Metcalfe, D.D. et al. Assessment of the allergenic potential of foods derived from genetically modified crop plants. *Crit. Rev. Food Sci. Nutr.* **36**(S), 165–186 (1996).
6. FAO/WHO. Evaluation of allergenicity of genetically modified foods. Report of a joint FAO/WHO expert consultation on allergenicity of foods derived from biotechnology. (Food and Agriculture Organization of the United Nations (FAO), Rome, 2001.
7. Goodman, R.E., Hefle, S.L., Taylor, S.L. & van Ree, R. Assessing genetically modified crops to minimize the risk of increased food allergy: a review. *Int. Arch. Allergy Immunol.* **137**, 153–166 (2005).
8. Aalberse, R.C. Structural biology of allergens. *J. Allergy Clin. Immunol.* **106**, 228–238 (2000).
9. Ladics, G.S., Bannon, G.A., Silvanovich, A. & Cressman, R.F. Comparison of conventional FASTA identity searches with the 80 amino acid sliding window FASTA search for elucidation of potential identities to known allergens. *Mol. Nutr. Food Res.* **51**, 985–998 (2007).
10. Goodman, R.E. & Hefle, S.L. Gaining perspective on the allergenicity assessment of genetically modified crops. *Expert Opin. Immunol.* **1**, 561–578 (2005).
11. Goodman, R.E. & Wise, J. Predicting the allergenicity of novel proteins in genetically modified organisms. in *Food Allergy* (eds. Maleki S.J. et al.) 219–247 (American Society of Microbiology Press, Washington, DC, 2006).
12. Hileman, R.E. et al. Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. *Int. Arch. Allergy Immunol.* **128**, 280–291 (2002).
13. Stadler, M.B. & Stadler, B.M. Allergenicity prediction by protein sequence. *FASEB J.* **17**, 1141–1143 (2003).
14. Silvanovich, A. *et al.* The value of short amino acid sequence matches for prediction of protein allergenicity. *Toxicol. Sci.* **90**, 252–258 (2006).
15. Sten, E. *et al.* A comparative study of the allergenic potency of wild-type and glyphosate-tolerant gene-modified soybean cultivars. *APMIS* **112**, 21–28 (2004).
16. Ladics, G.S., Bardina, L., Cressman, R.F., Mattsson, J.L. & Sampson, H.A. Lack of cross-reactivity between the *Bacillus thuringiensis* derived protein Cry1F in maize grain and dust mite Der p7 protein with human sera positive for Der p7-IgE. *Regul. Toxicol. Pharmacol.* **44**, 136–143 (2006).
17. van Ree, R. Carbohydrate epitopes and their relevance for the diagnosis and treatment of allergic disease. *Int. Arch. Allergy Immunol.* **129**, 189–197 (2002).
18. Altmann, F. The role of protein glycosylation in allergy. *Int. Arch. Allergy Immunol.* **142**, 99–115 (2007).
19. Asero, R. *et al.* IgE-mediated food allergy diagnosis: current status and new perspectives. *Mol. Nutr. Food Res.* **51**, 135–147 (2007).
20. Astwood, J.D., Leach, J.N. & Fuchs, R.L. Stability of food allergens to digestion in vitro. *Nat. Biotechnol.* **14**, 1269–1273 (1996).
21. Thomas, K. *et al.* A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regul. Toxicol. Pharmacol.* **39**, 87–98 (2004).
22. Asero, R. *et al.* Lipid transfer protein: a pan-allergen in plant-derived foods that is highly resistant to pepsin digestion. *Int. Arch. Allergy Immunol.* **122**, 20–32 (2000).
23. Fu, T.J., Abbott, U.R. & Hatzos, C. Digestibility of food allergens and nonallergenic proteins in simulated gastric fluid and simulated intestinal fluid—a comparative study. *J. Agric. Food Chem.* **50**, 7154–7160 (2002).
24. Yagami, T., Haishima, Y., Nakamura, A., Osuna, H. & Ikezawa, Z. Digestibility of allergens extracted from natural rubber latex and vegetable foods. *J. Allergy Clin. Immunol.* **106**,

752–762 (2000).
25. Aalberse, R.C. & Stapel, S.O. Structure of food allergens in relation to allergenicity. *Pediatr. Allergy Immunol.* **12** Suppl 14, 10–14 (2001).
26. Vassilopoulou, E. et al. Severe immediate allergic reactions to grapes: part of a lipid transfer protein-associated clinical syndrome. *Int. Arch. Allergy Immunol.* **143**, 92–102 (2007).
27. Pastorello, E.A. *et al.* Identification of hazelnut major allergens in sensitive patients with positive double-blind, placebo-controlled food challenge results. *J. Allergy Clin. Immunol.* **109**, 563–570 (2002).
28. Breiteneder, H. & Radauer, C. A classification of plant food allergens. *J. Allergy Clin. Immunol.* **114**, 127–130 (2004).
29. Untersmayr, E. *et al.* The effects of gastric digestion on codfish allergenicity. *J. Allergy Clin. Immunol.* **115**, 377–382 (2005).
30. Bannon, G.A. *et al.* Digestive stability in the context of assessing the potential allergenicity of food proteins. *Comments Toxicol.* **8**, 271–285 (2002).
31. Burks, A.W. & Fuchs, R.L. Assessment of the endogenous allergenicity in glyphosate-tolerant and commercial soybean varieties. *J. Allergy Clin. Immunol.* **96**, 1008–1010 (1995).
32. Lehrer, S.B. & Reese, G. Recombinant proteins in newly developed foods: identification of allergenic activity. *Int. Arch. Allergy Immunol.* **113**, 122–124 (1997).
33. Gao, Z.S. et al. Genomic cloning and linkage mapping of Mal d 1 (PR-10) gene family in apple (*Malus domestica*). *Theor. Appl. Genet.* **111**, 171–183 (2005).
34. Carnes, J., Ferrer, A. & Fernandez-Caldas, E. Allergenicity of 10 different apple varieties. *Ann. Allergy Asthma Immunol.* **96**, 564–570 (2006).
35. Zuidmeer, L. Allergenicity assessment of apple cultivars: hurdles in quantifying labile fruit allergens. *Int. Arch. Allergy Immunol.* **141**, 230–240 (2006).
36. Codina, R., Ardusso, L., Lockey, R.F., Crisci, C. & Medina, I. Allergenicity of varieties of soybean. *Allergy* **58**, 1293–1298 (2003).
37. Sancho, A.I. et al. Maturity and storage influence on the apple (*Malus domestica*) allergen Mal d 3, a nonspecific lipid transfer protein. *J. Agric. Food Chem.* **54**, 5098–5104 (2006a).
38. Sancho, A.I. *et al.* Effect of postharvest storage on the expression of the apple allergen Mal d 1. J. *Agric. Food Chem.* **54**, 5917–5923 (2006b).
39. Marzban, G. *et al.* Localization and distribution of the major allergens in apple fruits. *Plant Sci.* **169**, 387–394 (2005).
40. Jenkins, J.A., Griffiths-Jones, S., Shewry, P.R., Breiteneder, H. & Mills, E.N.C. Structural relatedness of plant food allergens with specific reference to cross-reactive allergens: an in silico analysis. *J. Allergy Clin. Immunol.* **115**, 163–170 (2005).
41. Ibanez, M.D., Marinez, M., Sanchez, J.J. & Fernandez-Caldas, E. Legume cross-reactivity (in Spanish). *Allergol. Immunopathol. (Madr.)* **31**, 151–161 (2003).
42. Bindslev-Jensen, C., *et al.* Assessment of the potential allergenicity of ice structuring protein type III HPLC 12 using the FAO/WHO 2001 decision tree for novel foods. *Food Chem. Toxicol.* **41**, 81–87 (2003).
43. Ferriera, F., Hawranek, T., Gruber, P., Wopfner, N. & Mari, A. Allergic cross-reactivity: from gene to the clinic. *Allergy* **59**, 243–267 (2004).
44. Dearman, R.J. & Kimber, I. A mouse model for food allergy using intraperitoneal sensitization. *Methods* **41**, 91–98 (2007).
45. McClain, S. & Bannon, G.A. Animal models of food allergy: opportunities and barriers. *Curr. Allergy Asthma Rep.* **6**, 141–144 (2006).
46. Matsuda, T., Matsubara, T. & Hino, S. Immunogenic and allergenic potentials of natural and recombinant innocuous proteins. *J. Biosci. Bioeng.* **101**, 203–211 (2006).
47. Perzanowski, M.S., Ronmark, E., Nold, B., Lundback, B. & Platts-Mills, T.A. Relevance of allergens from cats and dogs to asthma in the northernmost province of Sweden: schools as a major site of exposure. *J. Allergy Clin. Immunol.* **103**, 1018–1024 (1999).
48. Fernandez-Rivas, M. et al. Apple allergy across Europe: how allergen sensitization profiles determine the clinical expression of allergies to plant foods. *J. Allergy Clin. Immunol.* **118**, 481–488 (2006).
49. Reuter, A. *et al.* A critical assessment of allergen component-based in vitro diagnosis in cherry allergy across Europe. *Clin. Exp. Allergy* **36**, 815–823 (2006).
50. Herouet, C. *et al.* Safety evaluation of the phosphinothricin acetyltransferase proteins encoded by the pat and bar sequences that confer tolerance to glufosinate-ammonium herbicide in transgenic plants. *Regul. Toxicol. Pharmacol.* **41**, 134–149 (2005).
51. Jones, S.M., Magnolfi, C.F., Cooke, S.K. & Sampson, H.A. Immunologic cross-reactivity among cereal grains and grasses in children with food hypersensitivity. *J. Allergy Clin. Immunol.* **96**, 341–351 (1995).
52. Ikematsu, K., Tachimoto, H., Sugiasaki, C., Syukuya, A. & Ebisawa, M. Feature of food allergy developed during infancy (1)—relationship between infantile atopic dermatitis and food allergy (in Japanese). *Arerugi* **55**, 140–150 (2006).
53. Morton, R.L. *et al.* Bean alpha-amylase inhibitor 1 in transgenic peas (*Pisum sativum*) provides complete protection from pea weevil (*Bruchus pisorum*) under field conditions. *Proc. Natl. Acad. Sci. USA* **97**, 3820–3825 (2000).
54. Prescott, V.E. et al. Transgenic expression of bean alpha-amylase inhibitor in peas results in altered structure and immunogenicity. *J. Agric. Food Chem.* **53**, 9023–9030 (2005).
55. Editorial. Genetically modified mush. *Nat. Biotechnol.* **24**, 2 (2006).
56. Sinagawa-Garcia, S.R. *et al.* Safety assessment by in vitro digestibility and allergenicity of genetically modified maize with Amaranth 11S globulin. *J. Agric. Food Chem.* **52**, 2709–2714 (2004).

RESEARCH ARTICLE

AllergenOnline: A peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity

Richard E. Goodman[1], Motohiro Ebisawa[2], Fatima Ferreira[3], Hugh A. Sampson[4], Ronald van Ree[5], Stefan Vieths[6], Joseph L. Baumert[1], Barbara Bohle[7], Sreedevi Lalithambika[1], John Wise[1] and Steve L. Taylor[1]

[1] Food Allergy Research and Resource Program, Department of Food Science and Technology, University of Nebraska-Lincoln, Lincoln, NE, USA
[2] Department of Allergy, Sagamihara National Hospital, Sagamihara, Japan
[3] Department of Molecular Biology, University of Salzburg, Salzburg, Austria
[4] Department of Pediatrics, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[5] Departments of Experimental Immunology and of Otorhinolaryngology, Academic Medical Center, Amsterdam, The Netherlands
[6] Department of Allergology, Paul-Ehrlich-Institut, Langen, Germany
[7] Department of Pathophysiology and Allergy Research, Medical University of Vienna, Vienna, Austria

Received: October 1, 2016
Revised: February 4, 2016
Accepted: February 9, 2016

Scope: Increasingly regulators are demanding evaluation of potential allergenicity of foods prior to marketing. Primary risks are the transfer of allergens or potentially cross-reactive proteins into new foods. AllergenOnline was developed in 2005 as a peer-reviewed bioinformatics platform to evaluate risks of new dietary proteins in genetically modified organisms (GMO) and novel foods.

Methods and results: The process used to identify suspected allergens and evaluate the evidence of allergenicity was refined between 2010 and 2015. Candidate proteins are identified from the NCBI database using keyword searches, the WHO/IUIS nomenclature database and peer reviewed publications. Criteria to classify proteins as allergens are described. Characteristics of the protein, the source and human subjects, test methods and results are evaluated by our expert panel and archived. Food, inhalant, salivary, venom, and contact allergens are included. Users access allergen sequences through links to the NCBI database and relevant references are listed online. Version 16 includes 1956 sequences from 778 taxonomic-protein groups that are accepted with evidence of allergic serum IgE-binding and/or biological activity.

Conclusion: AllergenOnline provides a useful peer-reviewed tool for identifying the primary potential risks of allergy for GMOs and novel foods based on criteria described by the Codex Alimentarius Commission (2003).

Keywords:
Allergens / Bioinformatics / Food allergy / Genetically modified / Risk assessment

1 Introduction

Food allergy has become an important food safety issue that was hardly recognized 50 years ago. In the past 20 years the prevalence of IgE-mediated food allergy has grown from 1–2% overall to 5–8% in the United States, with similar numbers in the European Union [1,2]. Diverse environmental and behavioral factors likely contribute to the increased prevalence of sensitization and allergy, however, objectively defined causal factors to explain the increases are still lacking [3]. Control of manufacturing, sourcing, and labeling of ingredients in processed foods has become a major focus of regulators in many countries to help protect allergic consumers. However, this does not address risks associated with restaurant or home-cooked meals, which more frequently cause severe reactions [4]. While only a few foods and proteins are

Correspondence: Prof. Richard E. Goodman
E-mail: rgoodman2@unl.edu

Abbreviations: **AOL**, AllergenOnline.org; **FARRP**, Food Allergy Research and Resource Program; **FDA**, US Food and Drug Administration; **GM**, genetically modified; **GMO**, genetically modified organisms; **IgE**, Immunoglobulin E; **NCBI**, National Center for Biotechnology Information; **SPT**, Skin Prick Test

responsible for the majority of severe clinical reactions and the focus of food safety initiatives, many additional foods and proteins are "falsely" labeled as allergens in public sequence databases and scientific literature confounding regulatory dossiers.

The possibility that new food crop varieties developed through genetic engineering, often referred to as genetically modified (GM) crops, might cause an increased risk of food allergy was recognized in 1992 by the US Food and Drug Administration (FDA) (www.aphis.usda.gov/brs/fedregister/coordinated_framework). As described previously, development of a protective allergenicity assessment process based on scientific knowledge of allergy and allergens has been successful in protecting consumers [5–8]. To date there is no proof that an introduced protein in an approved GM crop has caused food allergy or risks for consumers. Arguably the most important assessment tool is the use of bioinformatics (sequence comparisons) to evaluate candidate GM proteins to prevent the transfer of an allergen or likely cross-reactive protein from one source into a different species used for food production [8].

2 Allergenonline.org database

The AllergenOnline.org (AOL) is maintained by the Food Allergy Research and Resource Program (FARRP) at the University of Nebraska-Lincoln. The database was developed in 2005 and provides a risk assessment tool for evaluating the potential allergenicity of new food proteins produced by genetically modified organisms (GMO) and novel protein ingredients in processed foods as well as a research aid. An overview of the management, expert panel, criteria for including proteins and financial support are described for transparency. The rationale for the processes of construction and updating the database are included. The protein risk assessment sequence search evaluation for GM products and novel proteins is described briefly but this manuscript focuses on the identification of allergenic proteins from sources causing allergic reactions upon ingestion, inhalation, contact, sting, and bite, respectively. Importantly there are no uniform objective characteristics that determine whether a protein will become an allergen. Investigators have tried to develop computer programs to predict the allergenic potential of proteins based on amino acid sequences or structural predictions, but so far none have proven predictive over diverse allergen types [9]. The most predictive comparison for risk evaluation is whether a new protein is at least 50% identical to a known allergenic protein based on local sequence alignment (FASTA or BLASTP) indicating that it may be cross-reactive in IgE binding or may be a bonafide allergen as observed by Aalberse [10]. Post-translational modifications including disulfide bond linkages, proteolysis, glycosylation, myristoylation, hydroxylation, or other alterations can change physical properties that might enhance or reduce potential allergenicity. The abundance of the protein, matrix composition, and the route of exposure may also impact potential sensitization or elicitation. The comparison may demonstrate whether a new protein is already an allergen or is sufficiently identical to an allergen to suspect likely cross-reactivity. It should be recognized that all allergens are not equal. Some are extremely potent while other proteins bind IgE in a high percentage of people, yet rarely, if ever cause reactions [11].

2.1 Defining allergens

The task of defining allergens is complex. Type I allergens are typically proteins that sensitize (induce both protein-specific Th2 helper cells and protein-specific IgE producing B cells) and elicit an IgE response by cross-linking IgE on surface receptors of mast cells and basophils, leading to mediator release and symptoms. However, nearly identical proteins may elicit a cross-reactive allergic response due to IgE binding to a structurally related protein. Allergic symptoms vary remarkably between subjects, and include urticaria, angioedema, asthma, hypotension, vomiting, diarrhea, or multisystem involvement defining anaphylaxis that can lead to death. Objective clinical reactions to individual proteins are rarely determined. Normally only the allergenic source is identified by testing with a complex mixture of proteins extracted from the source. While some publications suggest that allergens are typically glycoproteins of 10–70 kDa, the amino acid sequence, structure, size, abundance, stability in heat and to proteases and the sources of allergens are highly variable and few allergens fit all of those characteristics [5]. Demonstration of specific IgE binding to a protein in vitro does not prove that the protein will cause allergy in vivo as two or more IgE-binding sites are necessary to cross-link receptors on effector cells. Other factors including affinity of binding, abundance, and exposure will determine reactivity and relative risks. Identification of the protein(s) causing allergic reactions is complicated as most allergenic sources contain a number of proteins that bind IgE of some allergic subjects.

Thus deciding which proteins to include in a searchable database for risk assessment requires a rigorous review by allergy experts. We have set criteria to define *bonafide* allergens, putative allergens, or proteins with unproven allergenicity (Table 1). The latter comprise proteins considered as allergens based only on sequence homology, IgE binding without clinical disease, immune responses to parasites where IgE aids in protection, or sequences that are too short to cross-link IgE bound to mast cell or basophil FcεRI (Table 2).

2.2 Food allergy assessment of novel proteins for GM crops: USA and EU

In order to reduce health risks of GM crops, a consensus evaluation process was developed through consultations with a variety of scientists prior to 1990 in the United States [12]. The primary concerns for food safety of GM crops include

Table 1. Evaluating the evidence of allergenicity for new proteins

Evidence	Allergen	Putative allergen	Unproven
Evidence the protein is expressed in an allergenic source including the correct tissues	Direct protein by LC-MS/MS, N-terminal sequencing, specific antibody binding or as a minimum mRNA	Direct protein by LC-MS/MS, N-terminal sequencing, specific antibody binding or as a minimum mRNA	Negative
Description of source of protein as recombinant or natural source and characterization	Clear as to genus and species, appropriate organ-specific material (fruit, pollen, etc.), or cDNA with expression of recombinant protein	Recombinant protein from mRNA without proof the protein is in material humans are exposed to.	In adequate description or genomic sequence only
Allergic serum donors	Description of clinical symptoms for allergic subjects. Clinical challenge to the source and specific IgE binding to protein (e.g. blinded food allergen challenge, inhalation challenge, or skin prick tests)	Appropriate clinical description for allergic subjects. Specific IgE binding to an extract of the source, but no challenge with extract or natural source	Vague clinical symptom description (e.g. asthma, atopy) without reference to the source. Or serum IgE binding to protein without symptoms
In vitro IgE-binding data	Immunoblot or ELISA with pure protein and extract of source, with appropriate negative controls. Method described with appropriate controls. Inhibition of binding to protein using extract or potentially cross-reactive protein	In vitro IgE binding with recombinant protein only, without an extract or purified natural protein from relevant raw material	IgE binding to extract not pure protein, or IgE binding to protein without a benchmark of binding to extract or other control proteins, threshold set based on a very low threshold such as 2 Std. Dev. Above negative control
Protein includes potential cross-reactive carbohydrate determinants (CCD)	Demonstration that CCD is not responsible for binding, based on expression in *E. coli*, or proven deglycosylation	Presence of CCD not tested, probable glycoprotein from plant or arthropod, with IgE binding possibly due to CCD	
Proven biological activity of IgE binding	Skin prick test or basophil activation with pure protein with appropriate controls	Skin prick test or basophil activation without a benchmark of activation.	No biological assay and no clear specific IgE binding

The primary information used to categorize a protein as an allergen, or putative allergen is defined in peer-reviewed publications. Since few proteins meet all criteria, therefore a weight of evidence is used to judge the outcome. Proteins that do not meet minimum criteria are categorized as unproven but will be reevaluated if relevant new publications are identified.

the potential risks of allergenicity or toxicity from novel proteins and nutritional equivalence of the GM plant. Nearly 150 academic, regulatory and industrial scientists from various disciplines were involved in developing the FDA recommendations [8]. An important conclusion was that the safety of GMOs should be evaluated using existing U.S. food safety laws, administered by the FDA, Department of Agriculture (USDA), and the Environmental Protection Agency (EPA) [13].

Regulations and guidance documents in most other countries were developed following the 1992 US FDA recommendations. The utility of the assessment was demonstrated by the example of a potential new soybean product developed by Pioneer Hi-Bred, which incorporated a gene for the 2S albumin from Brazil nut to improve the nutritional quality of soybeans by increasing the cysteine and methionine content [6]. At that time no allergens from Brazil nut were identified as being responsible for eliciting allergic responses. However, because the gene was derived from a commonly allergenic source (tree nuts), in vitro serum IgE-binding tests were performed using sera from Brazil nut allergic volunteers. Additionally, skin prick tests (SPT) with extracts of Brazil nut, non-GM soybean, and the GM soybean were performed. Results demonstrated that the transferred gene encoded a previously unrecognized major allergen that would have put Brazil nut allergic consumers at risk if they consumed the GM soybean [6]. Therefore developers stopped development of that product, which was never used in food production. The

Table 2. Defining evidence for exclusion from the annual NCBI download for the expert panel review

Evidence	Explanation
Sequence Homology	NCBI entries labeled as "allergen" or "allergen-like" simply due to low-sequence identity matches
Human proteins	Although there are published reports of IgE binding to some human proteins, those are involved in autoimmune responses either due to high sequence identity to exogenous allergens (e.g. profilins) or to marked inflammation and auto-immune responses in those with lupus or arthritis.
Classification of protein type	Some entries are classified as allergens simply based on the class, or type of protein (e.g. trypsin inhibitor, SCP, MD2), without evidence of IgE binding or allergenicity
Patented Sequences	Individuals and companies are entering sequences in NCBI that are patented, including proteins that are defined as allergens without established proof of IgE binding or allergic elicitation; or sequences that might be useful for immunotherapy. These are not identifiable by the NCBI BLAST alignment tool. There are not links to information that would allow independent verification of allergenic activity.
Genome model organism sequences with "allergen-like designations"	Proteins from *Arabidopsis thaliana*, *Danio rerio*, *Drosophila melanogaster,* and other sequences from genome projects and without evidence of expression (e.g. mRNA, cDNA, or protein) are excluded unless there is a publication clearly linking the sequence to an allergy study.
Parasite proteins	Proteins from parasitic worms and arthropods are excluded unless there are clear published data demonstrating the source causes allergic disease in humans (not just in domestic animals), since the natural immune response to most parasites is generation of IgE antibodies and development of basophil an eosinophil responses. True allergy is elicited by only a few parasitic organisms, for example, the fish parasite *Anisakis simplex,* causes sensitization during first infection in humans with live worms, but can elicit allergic responses in later exposure to heat-killed worms in infected fish.
Sequences less than 8 amino acids long	The shortest useful sequences for bioinformatics inquiry regarding potential cross-reactivity are 8 or more amino acids long. Although unlikely to be unique in NCBI at 8 amino acids, there is no rationale for including sequences of less than eight.
No public reference of allergenicity beyond brief comments in NCBI or UNIPROT	Sequences that do not have associated publication references in the NCBI or UNIPROT databases and no identified reference in PubMed or Web of Science are not included in for panel review.

The following guideline was developed to eliminate thousands of protein sequences that have no evidence of allergenicity or are unlikely to have published evidence.

experience provided an important example of the predictive value of methods detailed by Metcalfe et al. (1996) and adopted by the CODEX Alimentarius Commission (2003), that serves as an international guideline for the safety assessment of genetically engineered crops [5, 7, 14]. Subsequently, the European Food Safety Authority (EFSA) used an extensive review process from 2007 until 2010 that produced recommendations for ways the assessment might be improved. The primary finding was the stated need for maintaining a curated, inclusive allergen database that is updated regularly [15].

2.3 Allergen protein databases for risk assessment

Individuals may be allergic to airway, contact or injected allergens and transfer of those allergens into a food source may cause food allergy. The risk assessment includes evaluate of the similarity of introduced proteins to all types of allergens. Researchers are identifying proteins as allergens from a variety of source organisms, but often the evidence of allergenicity is weak. In 1996, Astwood compiled a set of 219 allergens and gliadins (potential celiac eliciting proteins) from primary public sequence databases for risk assessment of Monsanto's GM products [16]. Prior to 2004 each biotechnology development company generated their own allergen databases for the evaluation of GM crops. Then seven companies agreed to fund an independent academic group, the FARRP at the University of Nebraska, to develop a common, publically available allergen database designed for use in bioinformatics searches to screen candidate GM proteins for potential risks of cross-reactivity [17]. The AllergenOnline.org database, (also referred to as the FARRP database), originally included protein sequences from the NCBI protein database that were labeled as "allergen" or "allergenic," as versions 5 (1189 total sequences) and version 6 (1537 total sequences) if expressed from a known allergenic source. However, a search for publications associated with many of the proteins in version 6 of AOL (unpublished, 2006) failed to identify scientific studies demonstrating human serum IgE binding or allergy to the source of the proteins. Therefore we developed and applied criteria for use in updates of AOL, version 7 that reduced the list of allergenic sequences to 1251 sequences. Since 2007 we have made minor changes in criteria yet man the number of sequences accepted as allergens has grown during annual updates as shown in the version history of AOL (Table 3).

While a few other allergen databases exist that might be used for risk assessment, they do not have a clearly

Table 3. Version history AOL sequence download and decisions

NCBI Searches for candidate allergens				AllergenOnline decisions (combined putative and bonafide allergens)			
Date range	Allerg*	Allergy	Allergen	Version #	Total # seq	Taxon protein groups	Species
1960/1/1–2005/06/01	33 992	16 934	4 376	**6**	1537	672	255
2005/06/01 – 2006/06/01	34 470	14 442	962	**7**	1251	451	221
2006/06/01 – 2007/06/01	8523	5211	2174	**8**	1313	483	229
2007/06/01 – 2008/06/01	27 558	24 754	2377	**9**	1386	502	236
2008/06/01 – 2009/06/01	126 837	123 259	3141	**10**	1471	529	254
2009/06/01 – 2010/06/01	138 033	133 234	3176	**11**	1491	553	265
2010/06/01 – 2011/06/01	227 716	223 351	3609	**12**	1603	603	273
2011/06/01 – 2012/06/01	308 529	299 853	3332	**13**	1630	612	275
2012/06/01 – 2013/06/01	935 229	921 032	8905	**14**	1706	645	290
2013/06/01 – 2014/06/01	932 310	911 991	19 793	**15**	1897	744	335
2014/06/01 – 2015/06/01	166 730	147 289	18 082	**16**	1956	778	345
NO date limit	**2 953 644**	**2 843 504**	**70 705**	**na**	**na**	**na**	**na**

Potential allergen sequences identified by keyword searches of the NCBI protein database using date limitations are listed as candidates along with the final determinations (cumulative) for versions of AOL. Note: for version 7, more stringent criteria were implemented reducing the number of allergens. The version number, total number of sequences, taxonomic protein groups and number of species accepted in AllergenOnline.org are shown (right side).

defined review process, or do not use the same risk assessment search algorithms. The WHO/IUIS Allergen Nomenclature Subcommittee maintains a database that provides recognizable names for specific allergens (e.g. Ara h 2 for the 2S albumin from *Arachis hypogaea*) based on taxonomy and protein types (www.allergen.org), prior to publication. The SDAP database has fewer sequences (1526) compared to AOL. Importantly, the review process has not been described and SDAP was last updated in February, 2013 (http://fermi.utmb.edu/SDAP/sdap_man.html). Proteins in the AllFam database (1091 allergens in 186 allergen families) are organized into structural protein families to provide information and associations to help identify potentially shared conformational epitopes. AllFam was last updated in 2011 (www.meduniwien.ac. At/allergens/allfam). The Allergome database (www.allergome.org) is a very good source of references for scientific information regarding allergy to sources of allergens and to many allergens. However, many of the entries in Allergome do not include sequences. The general sequence protein database, NCBI Protein (www.ncbi.nlm.nih.gov/protein) is a repository of more than 70 million protein (amino acid) sequences including those translated from GenBank, RefSeq, and TPA as well as sequences from SwissProt, PIR, PRF, and PDB as described on their website databases. The NCBI protein database is searchable by key word queries and provides a BLAST alignment tool for local sequence alignments to all entries, or to subsets based on user queries. However, keywords related to allergy are often entered based on auto-annotation from very low-level sequence identities to an allergen. Sequence searches of NCBI can be made using "allergen," but there are 69 000 plus "allergen" entries and most are not associated with published data demonstrating allergenicity. Thus the focus of the AOL is to provide a comprehensive, sequence searchable database of proteins based on publication of at least protein-specific IgE binding, and preferably biological activity using sera from subjects allergic to the source.

2.4 Allergenic protein sequence comparisons for food risk evaluation

Sequence searching algorithms that have been recommended for evaluating food allergy risks of GM proteins use a full-length FASTA (or BLASP), a sliding 80mer FASTA and/or exact 8 amino acid matching algorithms to compare the newly introduced protein with sequences of allergens to identify proteins that allergens or are likely to pose a risk of cross-reactivity [9, 18, 19]. The predictive value of the computer search depends on the quality of the database (inclusion of allergens, exclusion of proteins without proof of IgE binding or elicitation of allergy), the specificity of search algorithms and the criteria of significance. Although some researchers claim to have developed computer programs that predict B-cell epitopes, potential IgE binding is predicted by estimating structural motifs based on amino acid properties including charge hydrophobicity and nearest neighbor analysis, 3D surface plots or similarities to other known allergenic epitopes [20–22], but without broad validation. Development and maintenance of allergy to a given protein requires establishment of both T-helper cell epitopes and B-cell epitopes. However, specific proteins may elicit reactions through cross-reactive recognition of IgE epitopes without T-cell help [10]. The great diversity in individual allergic patient responses

to different proteins in complex food and environmental allergenic sources and of epitope recognition to the same protein makes validation of predictions complex as demonstrated by recent studies of allergen microarrays for allergy diagnosis [23, 24]. Based on current studies it seems that the best method to predict potential risks of allergy or cross-reactivity is still to use a local alignment method (BLASTP or FASTA) with identity scores of greater than 50% to identify proteins that are known to be allergens or so similar in sequence to indicate a need to perform serum testing using samples from those with allergy to the matched protein as recommended by the CODEX Alimentarius Commission in 2003 [9, 14]. The CODEX Alimentarius Commission guideline is not a regulation, however, CODEX member countries including members of the European Union are expected to follow CODEX [7, 25, 26]. The guidance suggests "significant sequence identity matches" are those exceeding 35% identity in alignments over segments of 80 amino acids or longer. In addition, short segments (six or eight) amino acids with 100% identity match were suggested as possibly being meaningful although the predictive value of such matches without overall long identity alignments has been largely discredited [9, 17]. As scientific knowledge is gained about allergy and allergens, new methods and criteria may eventually be accepted if proven predictive [27]. Two studies suggest using an alternative to the sliding window of 80 amino acid approach, by using the statistical *E* score from FASTA or BLASTP to judge the relevance of overall identity matches might help in judging potential cross-reactivity when used to evaluate the significance of percent identity scores [28, 29]. However, *E* scores vary markedly with the size of the database and the length of aligned proteins. The *E* scores should be more useful if utilized as an additional criterion for judging marginal identity alignments (e.g. 35–50%) rather than being used as the primary criteria for potential cross-reactivity [17]. The results of the bioinformatics search inform developers and regulators regarding whether human serum IgE studies are warranted. However, appropriate serum testing requires well-characterized allergic and control sera, specific control and test substance characterization, validation of IgE detection reagents and inhibition to verify specificity of results [6, 9, 30, 31].

3 Origin of allergenonline database and selection of candidate allergens

The first public version of the AOL was compiled in 2005 (called version 5). It included 1189 protein amino acid (aa) sequences that represented 608 protein-taxonomic groups including isoallergens and variants (isoforms) as defined by King et al. [32]; from a total of 208 species of plants, animals, and fungi [16]. That dataset was seeded from Monsanto's 2004 database, which is an expanded version of the ALLERGEN3 dataset described by Hileman et al. [33]. Version 5 of AOL also included the entries from four other biotechnology developers (Bayer CropScience, Dow AgroSciences, DuPont-Pioneer, and Syngenta AG). Duplicates and substrings of longer sequences were removed. Additional sequences were added through keyword searches (allerg*) of the NCBI Protein database.

Sequences in the AOL include environmental allergens representing proteins from pollen of flowering plants, common molds, house dust mites (HDM), and insects; contact allergens (e.g. latex) and food allergens. The complete dataset of unique sequences (>2100 proteins) was compared to NCBI sequence entries (February, 2005) to verify annotation and the appropriate GenInfor Identifer (GI) number During annual AOL updates (2006–2015) large increases in the number of keyword identified sequences were identified without proof of allergy as shown in Table 3. Table 3 also shows the number of sequences, groups, and species that were accepted as allergens based on criteria described below. The sequences are organized in taxonomic-protein groups based on unique species (or occasionally two species) and protein type (at least 65% identical, similar functional protein).

Because literature searches and reviews are quite time consuming, we developed a screening system for removing a large number of proteins that are highly unlikely to cause IgE-mediated allergies (described in Section 2.2).

3.1 Identifying new candidate allergenic proteins

Many well-characterized allergens have been identified from clearly defined allergenic sources over the past few decades. Understanding the source and exposure are important considerations for inclusion. Food allergens are often well characterized as exposure is usually identifiable and controllable, allowing clear definition of study subjects. While diverse foods may occasionally cause allergy, eight foods, or food groups (peanuts, soybeans, tree nuts, milk, eggs, fish, crustaceans, and wheat) are thought to account for nearly 90% of reactions in the United States [5, 34]. Recent studies demonstrate that even among the common allergens, peanut, and a few tree nut species (e.g. almond, walnut, pecan, and hazelnut), shrimp and milk cause most of the fatal and life-threatening reactions [35]. Only a few proteins in each of these complex food sources bind the majority of IgE and are likely responsible for eliciting serious reactions [36, 37]. But identification of the biologically important allergens is difficult. While a large percent of people believe they have food allergy, careful diagnosis using IgE binding, skin prick tests, and food challenges demonstrate no food allergy, even if individuals have IgE that binds to a protein of the source [38]. A recent study of 4000 cases of severe anaphylaxis in central Europe implicated many foods commonly causing anaphylaxis in the United States (peanuts, tree nuts, milk, and crustaceans, fish, and hens egg) although a surprising number of wheat and celeriac allergic subjects were also identified [39]. Population studies using the same reagents and methods in China, India, and Russia demonstrated much higher rates of sensitization (IgE binding) than allergy as judged by lack of SPT reactions

and histories [40]. However, allergy in young children to commonly consumed foods, e.g. milk and egg, were similar to those in the United States and EU [40]. A new study in the EU demonstrated that self-reported allergy and specific IgE binding using an allergen-protein microarray overestimated clinical reactivity [41].

The dominant allergenic proteins in peanuts in various populations are 2S albumins (Ara h 2 and Ara h 6), 7S vicilin (Ara h 1), and 11S glycinin (Ara h 3) according to Palmer et al. [42] and Vickery et al. [43]. Homologous proteins in walnut have been identified as the major allergens in walnut [44]. Yet in spite of moderately high sequence identities (33–53%) between peanut and walnut homologous allergens, there is little evidence that they are cross-reactive [44]. A nonspecific lipid transfer protein (nsLTP) is the major allergen identified in peach, which can cause severe food-allergic reactions in some subjects from Italy and Spain [45, 46], but similar allergies are rarely reported elsewhere. While nsLTP sequences and structures are highly conserved in fruits of *Rosaceae* and across much broader taxonomic groups and in vitro IgE binding is often shared, allergy to the sources is often much more restricted [47, 48]. Many other proteins from these foods have been identified as allergens. While careful studies have demonstrated IgE binding to a number of other proteins from these sources, the potency, and relevance of many for allergy is hard to demonstrate.

Inhalation allergy and allergens are often difficult to diagnose and proof of the identity of the proteins responsible for eliciting allergic reactions is often lacking. The abundance and time of exposure is important. There is little doubt that the pathogenesis related proteins (e.g. PR-10) of birch pollen (Bet v 1) and related trees are prominent allergens for many subjects experiencing rhinitis, conjunctivitis, and asthma in the spring in northern Europe [49, 50]. However, there is less proof that the profilin of birch pollen (Bet v 2) or cyclophilin (Bet v 7) are important allergens. Conservation of the amino acid sequence and three dimensional structures of proteins across diverse taxonomic organisms (birch, apple, carrot, and hazelnut) means that sera from some donors allergic to birch pollen may bind to homologues in apple, carrot, and hazelnut by in vitro IgE assays. More than 25 proteins from two species of *Dermatophagoides sp*. are listed as allergens in the WHO/IUIS allergen database, however, few of them have been proven to be significant allergens [51]. Further complexity is added due to the existence of other diverse species of mites and arthropods that are from other environments, with different possibilities of exposure for the individual.

Studies reporting proteins as allergens often rely on in vitro protein-specific serum IgE-binding data. The formats of the tests vary from simple immunoblotting of extracts followed by purification of identification of proteins by MS or N-terminal sequence to IgE binding to highly purified natural or recombinant proteins purified and presented in native or denatured forms. Typically inhibition assays with homologous proteins from other sources are required to demonstrate specificity of binding. Appropriate negative control sera and nonallergenic proteins are required to demonstrate specificity. The specificity of secondary reagents, blocking methods should be well described as reviewed and demonstrated in a number of studies [30, 52]. Test proteins must be well characterized, appropriately folded, and presented to provide accurate IgE binding compared to the multiple proteins present in the source materials in order to appropriately diagnose allergy and identify allergens [53–55]. However, IgE binding to common pan-allergens may be misleading without inhibition data [47, 52].

3.2 Excluding proteins that are unlikely to represent risks of allergy

In order to reduce the number of likely irrelevant proteins that are not from allergenic sources and may be not be expressed, we have developed a set of criteria to exclude sequences from NCBI that are highly unlikely to be *bonafide* allergens (Table 2). However, annotation of NCBI sequences changes over time. Thus each year we challenge the screening process to reduce the likelihood of missing potential allergens. The rapid increase in the number and diversity of genomes that have been sequenced recently and annotated with predicted (hypothetical) proteins labeled as allergens or allergen-like proteins is surprising. Data fields of NCBI entries have to be screened to exclude entries where the only association with allergy is that the authors are from an allergy institute (e.g. National Institute of Allergy and Infectious Diseases) or the organism is known to cause allergies (e.g. Aspergillus). Criteria defining major groups for exclusion are listed in Table 2. Candidate allergen entries and information from publications describing studies of allergenicity related to the taxonomic source or the specific proteins are entered in a private MySQL database in FARRP to provide a mechanism for reviewing and archiving for each annual review. Detailed descriptions, expert opinions, referenced publications, and decisions are recorded and maintained by FARRP each year. The public database (http://www.allergenonline.org/) only shows sequences of proteins with sufficient published evidence of allergy, at a minimum-specific IgE binding from sera of subjects allergic to the source, to suggest a role in elicitation of symptomatic allergy. The public AOL database includes a description of the database (Home, About, and Contact) with a link to the version's historical record. The Browse function shows all accepted allergens, with species name, IUIS allergen name (if available), GI: number links to protein sequences and a links on the Group column listing relevant references. The Sequence Search link goes to a sequence entry box (FASTA format or plan text sequence, with or without numbers) and has a link to a Support page that describes the full-length and 80mer FASTA searches as well as an exact eight amino acid identity match section with suggestions for interpretation.

The download search criteria and resulting total candidate sequence numbers from the NCBI search for the current

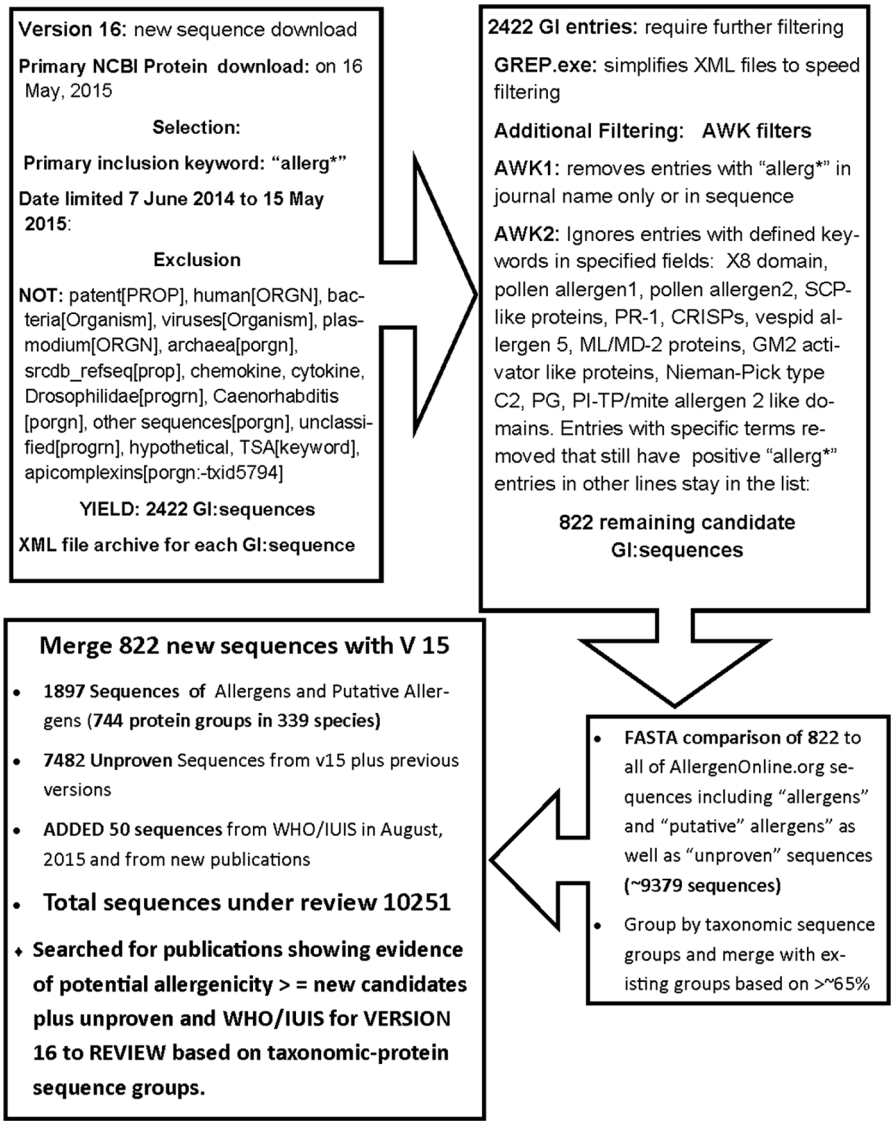


Figure 1. Download process of new candidate protein sequences for version 15. The upper-left box lists the query terms for NCBI to accumulate candidates. A total of 1511 sequences were gathered and subjected to further filtering with AWK filters to remove irrelevant entries (upper-right box), bringing the list to 558 candidates that were compared to all preexisting entries (>9100 sequences) in the database archive (lower right box) using the FASTA3 comparison. All new entries (in existing groups or new groups) were reviewed by the expert panel. In addition, the "unproven" groups were reviewed for new published information and reviewed by the panel if there was new information.

version (v16) are shown in the upper left box of Fig. 1. The search was downloaded from the NCBI Protein database on 15 May, 2015 as the end date using "allerg*" as the inclusive primary keyword in the Entrez Query, followed by terms to limit the entries based on date (beginning June 7, 2014) and other group identifiers as explained here. Exclusion of major groups of sequences that are highly unlikely to include allergenic proteins was accomplished using the Boolean logic terms "NOT" and "AND" in the search query string with keywords and categories (Fig. 1). Patented sequences are excluded because of limited information to judge their involvement in allergy. A number of patented protein sequences are modified for reduced risk of IgE binding for therapeutic use. Human sequences were excluded as we consider human proteins not to be allergens, although some display high sequence identity to allergens of molds, and some act as targets of auto-immune reactions [55, 56]. However, there is no evidence that human proteins are the primary sensitizers. Sequences from common genomic model organisms (e.g. *Arabidopsis thaliana*, *Danio rerio*, *Drosophila melanogaster*) are excluded as these organisms have not been demonstrated to cause allergies in humans. However, they contain many sequences labeled as allergenic or allergen-like, based on sequence homology. Viral sequences are excluded as there is no evidence that they elicit allergies. Specific viruses certainly skew immune responses, e.g. increase or exacerbate atopic asthma by enhancing the expression of signals including thymic stromal lymphopoietin (TSLP) or interleukins-25 and -33 [57, 58]. Respiratory syncytial viral (RSV) has been shown to skew the immune response away from a Th2- allergic response even though RSV and human rhinovirus are known to lead to development of asthma [59, 60]. While a number of mechanistic studies point to various immunomodulatory signals or outcomes induced by viruses, the evidence is primarily from studies in animal models. Evidence suggesting IgE binding to some viral proteins implicates cross-reactivity

that is unlikely to elicit reactions as demonstrated for IgE binding to VP4 of rotavirus [61]. Another investigation claims that a yellow fever virus vaccine caused anaphylaxis in 40 individuals (out five million recipients) to the first injection. However, it has been suggested that these reactions were not due to viral proteins, but rather egg proteins or gelatin used as carrier proteins [63]. Similarly, sequences from plasmodia, bacteria, and whole organism genome sequencing projects and immunological marker proteins (cytokines and chemokines) are excluded from the broad NCBI download using the search scheme shown in Fig. 1. This screening and filtering process allows us to avoid having to perform PubMed searches on thousands of proteins that are from organisms without known histories of allergy. Proteins from parasites are complicated to evaluate for potential allergenicity because in a number of cases, IgE and Th2 response provide immune-mediated protection against a number of parasites due to IgE triggering mast cells and helping develop eosinophil responses [64, 65]. However, there are a few parasites that clearly cause allergic responses. Humans can be infected with *Anisakis simplex* if ingesting raw, parasite infected fish. After the parasite dies in humans, some experience parasite protein-specific allergy when consuming well-cooked fish with dead *Anisakis simplex* worms [66]. *Ascaris sp*. has been suggested as another source of allergens. Therefore we search for publications of specific parasite proteins found in the WHO/IUIS database (www.allergen.org) and referenced in PubMed.

3.3 Integrating new entries with existing entries

The version 16 download identified 2422 potential new allergen sequences from NCBI (Fig. 1). A second set of screening filters (AWK filters) was used to filter out additional sequences that were annotated simply based on similarities to a few protein types (e.g. GM2, pollen allergen 2) without evidence the proteins have been detected or studied, leaving 822 sequences to integrate into existing entries (1897 allergens and 7482 unproven sequences). An additional 50 candidate entries were later identified (August, 2015) by a comparison to the WHO/IUIS database (www.allergen.org) and from our review of publications. The entries were compared to previous entries in AOL using FASTA to organize the entries in previously defined or new taxonomic-protein groups. The groups are defined by having 65% identity or more to the previous clearly defined allergen or within a new group. Specific PubMed literature searches were used to identify bacterial and viral proteins that might have been identified as allergens based on direct studies, but were missed due to our exclusion criteria since a few sequences from bacterial or viral sources may stimulate an allergic response [67–69].

Specific literature searches were performed for groups that were previously identified as "unproven" as well as new candidates. Many references were identified from the sequence entries in NCBI although those often describe cloning and characterization of the protein and may not have relevant information regarding proof of allergy. Keyword searches of PubMed looks for references to the organism, the authors of sequences and the protein. Complex searches are needed for some searches to uncover relevant papers or to limit publications using Boolean terms. Publications were also identified in the WHO/IUIS database and occasionally Allergome. Literature searches were performed by two people (SL and RG). Copies of the articles deemed relevant are loaded in the review database for RG and other members of the expert panel.

4 Criteria for allergens, allergic subjects, and test interpretation

From 2005 to 2007 the expert panel developed criteria for selection of allergens, taking into consideration the uncertainties of accurate diagnosis of allergy, methods for characterizing proteins, serum testing methods and biological assays that have been used to identify clearly allergenic proteins (e.g. 2S albumins, lipid transfer proteins, vicilins, tropomyosins, proteases of house dust mites) for classifying proteins as proven allergens, putative allergens, or unproven (incomplete or absence of evidence of allergenicity).

4.1 Source of the protein

The protein must be isolated from natural sources or synthesized from cDNA of a source of proven allergy. Characterization must include confirmation of at least a partial protein sequence that matches a translated sequence of cDNA or genomic clone, or a nearly full-length amino acid sequence from purified protein with a description of methods of purification and characterization. Immunoblots of extracts associated with human exposure are useful in addition to IgE binding to purified protein. Recombinant vectors and hosts must be appropriate for producing recombinant proteins depending on the characteristics of the protein; considering glycosylation, disulfide bond formation, or other probable posttranslational processing. If the complete amino acid sequence of the protein tested for IgE binding is listed in the publication, the GI: number and characterization (including partial amino acid sequence) must be described.

4.2 Serum IgE binding

Natural, purified proteins, or recombinant proteins may be used as the target for serum IgE binding. Test methods must be described and may include immunoblotting from SDS-PAGE under reducing or nonreducing conditions; ELISA, radioallergosorbent assay (RAST), dot-blotting with purified protein, or microarray assays with a description of test and control materials. Irrelevant or nonallergenic control proteins should be used in the same assay along with nonallergic sera

to demonstrate that any secondary detection system (e.g. anti-IgE) is specific. The source of specific critical reagents used in the assays should be clearly described (e.g. monoclonal or polyclonal anti-IgE; company and product number).

Description of serum donors used for tests must be included. Symptoms should be described and associated with the source of the protein. The use of "atopic" sera without an association of symptoms to the source by the subject is not acceptable. Allergies to foods are often clear if the source is consumed intact (tree nuts or fruits). However, many sources of food allergy (e.g. corn, soybean, wheat) are typically found in processed foods that contain multiple ingredients. Identification of the causative agent often requires confirmation by serum IgE tests, skin prick tests, elimination diets, or even food challenges [70]. Symptoms for food allergy usually begin within 10–20 min following ingestion [70,71]. However, some subjects experience delayed reactions four to eight hours after consumption, as recently described for allergy to red meat in individuals sensitized to the galactose-α-1,3-galactose glycan on nonhuman mammalian muscle proteins, mimicking epitopes of tick salivary proteins [72]. Airway allergy including rhinosinusitis or allergic asthma may be seasonal due to pollen exposure or mold exposure during peak seasons, or perennial due to pet dander, house dust mites, molds or cockroach in indoor environments. Diagnosis is typically based on history and skin prick tests with appropriate extracts or in vitro IgE-binding assays. Venoms from stinging and biting insects, spiders, or reptiles are often easily diagnosed, but are generally not of common occurrence. Self-reported symptoms by the study subjects without specific IgE binding or other antigen-specific tests are questioned and in most cases not accepted. Studies reporting results with multiple subjects having allergy to the same source are much more convincing than single case reports, but the expert panel recognizes that the prevalence of allergy to many allergenic sources is low. Well-conducted single case reports can be more convincing than studies using a number of poorly characterized subjects. Scoring information (e.g. SPT, wheal diameter; ImmunoCAP reagents and scores) and positive and negative control reagents should be described. A clear demonstration that a well-characterized protein specifically binds IgE from appropriately allergic human subjects is taken as proof that a protein is at least a putative allergen.

Studies reporting allergic reactions in nonhuman animals are not accepted as demonstration that a given protein is an allergen for humans. Therefore proteins that have only been tested in rodents or domestic animals are excluded since the intended purpose of AllergenOnline.org is evaluating food safety for humans.

4.3 Biological activity

IgE antibody binding alone is not sufficient to merit classification as a *bonafide* allergen [73]. Allergic responses require effective cross-linking of IgE bound to FcεRI on the surface of mast cells or basophils to cause release of histamine and other mediators. Therefore, the ultimate proof that a protein is an allergen is through biological activity, demonstrated by skin prick tests with pure protein, specific challenges (airway challenge for inhalation allergens or food challenge for dietary allergens) or by in vitro basophil activation or histamine release [73,74]. Our consensus opinion is that demonstration of basophil activation or skin prick test activity to the purified protein demonstrates sufficient evidence to consider a protein as a *bonafide* allergen.

Proteins that are highly similar in sequence and from the same taxonomic group to an accepted allergen, according to our criteria, should not require the same level of testing and evaluation as a new, unique candidate. Therefore proteins from a single species sharing more than 65% identity across most of the sequence are accepted as equivalent.

Since version 7 there have been only a few minor changes in our criteria, which are reviewed every year by the expert panel. Since few publications match our ideal data recommendations, some flexibility is incorporated in the system. For instance, we prefer studies testing at least ten allergic subjects and two to ten negative controls (those not allergic to the material under investigation). However, since specific allergies to some sources are relatively rare, exceptions using one to five allergic subjects may be accepted. Negative controls should include subjects with allergies to unrelated sources who are exposed to the allergenic source material without reacting.

4.4 Special exclusions

A few glycoproteins have been identified to bind IgE from atopic subjects selectively to asparagine linked, complex carbohydrate determinants. Some research indicates that such glycoproteins are unlikely to elicit allergic responses following their ingestion [75]. For example, the nonglycosylated recombinant tomato protein Lyc e 2 (also called Sola l 2) produced in *E. coli* did not bind IgE from tomato-allergic subjects. The native glycosylated protein caused basophil histamine release whereas the recombinant protein without glycan did not [76]. The expert panel will reevaluate the rationale for keeping this protein in AllergenOnline during the current update for version 17. A similar challenge is presented by the celery allergen Api g 5 [77]. Those two proteins are in the current version of AOL.

5 Review process

After new candidate proteins are sorted into taxonomic protein groups and references are loaded in the review database, if we did not find relevant publications, the database manager (RG) searches again during his review of all entries. Text boxes are used to describe the review and findings. For groups voted as unproven in one year, the group is re-reviewed again

for over the next two annual updates. Each group is then assigned to two additional experts for review and recording their comments. Grading decisions are recorded by selection in a drop-down menu as "allergen," "putative," or "unproven." The reviews are then locked down and cannot be changed except by the database manager. When groups are fully scored, a final voting system is prepared and the reviewers have to revote, with the previous scores visible. They can check the detailed records prior to voting. The database manager then reviews all votes. If there is more than one dissenting vote or the arguments are compelling for a different outcome than voted by the majority, the database manager summarizes the dispute and returns the entry to the other experts for a rereview. Approximately one month after completion of final votes, the sequences and GI: numbers are checked again to ensure accuracy compared to the NCBI entries. In version 16 there are three entries that are not in the NCBI Protein database, but have been manually entered into AOL. Taxonomic-protein groups judged as "allergen" or "putative allergen" are referred to as allergens in the public database. Groups judged as "unproven" are held in the review system for later reevaluation in the next review cycles.

An archive of the new version is submitted to the database sponsors approximately one month prior to public release so they can prepare the databases (behind their firewall) that they use for regulatory submissions. A PDF of the complete set of allergens is included generated and posted on the AllergenOnline public website for convenience of those preparing regulatory dossiers.

5.1 AllergenOnline example decisions on candidate proteins

A sample of entries, references, and decisions are shown in Table 4. Five example proteins from diverse source organisms (cashew, *Anisakis sp.* parasite, honey bee venom, and house dust mite) judged as having clear evidence of allergenicity are described. Three examples judged as putative allergens (kiwi protein; *Alternaria sp.* mold protein and dog lipocalin) are described. And two sequences judged as unproven (a very short peptide from a protein of black gram, *Vigna mungo*; and a protein from Chinese jujube) are described.

5.2 Bioinformatics methods for safety evaluation of a novel protein

A hypothetical example of a protein that might be used in a GM crop {2S albumin (193 AA) in *Eucalyptus grandis* (GI:702384944)} was compared to AOL version 16. The full-length FASTA identified the best overall alignments with the 2S albumin of black walnut (*Juglans regia*) based on the smallest *E* score (8.9e-006) and 37.6% identity over a 181 AA alignment, which indicates probable homology. By a sliding 80 amino acid window, FASTA identified highly significant alignments to 2S albumins of hazelnut (48.7%), pecan (45%), English walnut and black walnut (45%), sesame (42.5%), and castor bean (41.2%). Only a single 8 AA identity match was found using the 8mer identity match, to the castor bean 2S albumin.

6 AllergenOnline expert panel

Individual experts who have participated on the expert review panel are listed on the website and are included as coauthors. All have significant extensive records of publishing on allergy and allergens. We maintain records of decisions by the panel and provide the sponsors with an archival copy of each new version of the database (www.AllergenOnline.org) that they subscribe to.

7 Public access to the AllergenOnline database

The FARRP at the University of Nebraska maintains the database as a free public service database. We do not track users or record or view any sequences entered by users for searches.

8 Discussion

The AOL was established in 2005 to provide a tool for evaluating the safety of dietary proteins that might be introduced into foods through genetic modification of plants, animal, or microbes and used in food production, or proteins that might be introduced as novel ingredients through food processing. It is intended to provide a sequence and information dataset representing known and putative allergenic proteins of foods and environmental sources (inhalation, contact, or injected). Selected sequences represent known proteins from allergenic sources, which specifically bind IgE from some individuals allergic to the source. The database is not intended to include every minor variant of sequences of every allergen, as publications and available comprehensive databases include many redundant and theoretical sequences. References used to judge taxonomic-protein groups are listed, but cannot be provided to the public due to copyright laws. Novel proteins that match proteins from AllergenOnline at a level above the conservative threshold of identity recommended by regulatory agencies should be considered for testing to evaluate specific IgE binding and allergy if the novel protein will be present in food.

It must be recognized that the level of proof and the potency of allergenicity of the proteins in sequences selected in the database are not equal. Some are clearly proven, abundant, and potent allergens from allergenic sources demonstrated to cause systemic anaphylaxis in sensitized individuals. There is evidence that many others are likely to contribute to the aller-

Table 4. Examples of allergens, putative allergens, and unproven, version 16

Genus/species	Candidate proteins	GI# protein	Publications	Current status
Anacardium occidentale Cashew	Legumin seed storage protein WHO/IUIS Ana o 2.0101	25991543	[79] Wang et al., 2003	Allergen
Anisakis simplex Fish parasite	Cysteine-protease inhibitor WHO/IUIS Ani s 4.0101	47605398 110346534	[80] Moneo et al, 2005, [81] Rodriguez-Mahillo et al., 2007	Allergen
Apis mellifera Honey bee	Dipeptidyl peptidase venom protein WHO/IUIS Api m 5.0101	187281543	[82] Blank et al., 2010	Allergen
Betula verrucosa Betula pendula Birch tree pollen	Pathogenesis related protein 10 (PR10) WHO/IUIS Bet v 1.0101	17938 Many more	[49] Swoboda et al., 1995 [50] Erdmann et al., 2005	Allergen
Dermatophagoides pteronyssinus European house dust mite	ML domain lipid binding protein WHO/IUIS Der p 2.0101	9280543 Many more	[83] Chua et al.,1990 [84] An et al., 2013	Allergen
Alternaria alternata Plant pathogenic fungus	Manganese superoxide dismutase WHO/IUIS Alt a 14.0101	529279957	[85] Postigo et al., 2011 [86] Gabriel et al., 2015	Putative allergen
Canis lupus familiaris Dog dander lipocalin	Lipocalin like protein WHO/IUIS Can f 4.0101	262232390	[87] Mattsson et al., 2010 [88] Niemi et al., 2014	Putative allergen
Actinidia deliciosa Kiwi fruit	Pectinmethylesterase inhibitor WHO/IUIS Act d 6.0101	27544452	[89] Irifune et al., 2004 [75] Bublin et al., 2011	Unproven
Vigna mungo Black gram	28 kDa protein purified by only 12 amino acids determined	410591582	[90] Kumari et al., 2012	Unproven
Ziziphus jujube Chinese jujube tree	Bet v 1 homologue Erroneous IUIS-like number	631901704	No publication	Unproven

gic response. Yet for many sequences included in AOL, there is only evidence of weak IgE binding or the proteins may represent minor components in suspected allergenic sources. Interpretation of results obtained from the use of the AllergenOnline database is the responsibility of the user. High percent identity matches to sequences with many clear references presenting objective symptoms associated with allergic disease indicate probable risk for some allergic consumers. Modest or marginal identity matches probably represent low risk. High identity matches to proteins from sources rarely reported as causing allergy or only subjective symptoms probably pose little risk. However, the context of future exposure and processing of foods that might include those proteins should be taken into consideration.

During the past 20 years, the number of putative allergens has increased rapidly. Yet less than 2000 proteins have been demonstrated to specifically bind IgE from subjects experiencing allergic symptoms following exposure to the respective proteins or protein source. Even less of those proteins have been demonstrated to have specific biological allergenic activity defined by SPT or basophil activation. The majority of data is from studies in Europe or North America. However, allergy studies are now being reported from diverse geographies where different source organisms of food and environmental organisms exist and the list of diverse allergens is expanding. The risk assessment process has become more challenging during the past 10 years as the genomes of many organisms are being added to the database accompanied by auto-annotation of proteins as "allergens" or "allergen-like" based on low level sequence or structural similarity.

The AOL is intended to help protect consumers around the world by providing a tool that is updated regularly based on global scientific information to help developers and regulators protect allergic consumers from unexpected and accidental exposure to their allergens in new food sources. The AOL has been used by the sponsoring companies to evaluate potential risks of allergenicity of their GM proteins. It has also been used in studies to evaluate potential products developed by nonprofit organizations and academic developers for example, Indian mustard plants designed to enhance hybrid seed production [78].

Financial support for the AOL was provided (2004–2015) by contractual agreements between the University of Nebraska-Lincoln (UNL) and the following companies: BASF Plant Science, Bayer CropScience, Dow AgroSciences, DuPont-Pioneer, KWS SAAT, Limagrain (CIE & Vilmorin), Monsanto Company, Proctor and Gamble (one of the seven original sponsors); as well as the Food Allergy Research and Resource Program at UNL. BASF, DuPont-Pioneer, and FARRP faculty are continuing to support AOL in 2016 and beyond. We gratefully acknowledge Sue Hefle (deceased) who was a Co-Director of FARRP and played an important part in development of the concept of the collaborative database.

R.G. drafted the manuscript, but all authors reviewed and contributed. R.G. and S.T. are the project leaders. J.W. was responsible for developing the web-based review system and the public database website, which is now maintained by S.L. The expert panel (S.T., J.B., H.S., M.E., R.vR., B.B., F.F., R.G., and previously S.V.) develop the guidelines for criteria used to classify proteins as allergens, putative allergens, or unproven proteins and provide the technical allergy review. R.G. is the database manager with overall responsibility for ensuring the reviews and records are complete.

Contract support for the allergen database sponsors paid partial salaries for R.G., J.W., and S.L. The external panelists (B.B., F.F., H.S., M.E., S.V., R.vR.) receive an honorarium and expenses for attending the annual expert panel meeting. S.V. has received grant funding previously for scientific studies from Monsanto and Pioneer. R.G. has received grant funding for previous scientific studies from BASF, Bayer, Monsanto, and Pioneer. Expert panel members are not assigned as primary or secondary reviewers of publications or entries they coauthored. Authors are screened by R.G. during triage. Reviews are recorded for each expert and archived. All experts perform final votes on all entries, which are recorded.

The authors have declared no conflict of interest.

9 References

[1] Sicherer, S. H., Epidemiology of food allergy. *J. Allergy Clin. Immunol.* 2011, *127*, 594–602.

[2] Rona, R. J., Keil, T., Summers, C., Gislason, D. et al., The prevalence of food allergy: a meta-analysis. *J. Allergy Clin. Immunol.* 2007, *120*, 638–646.

[3] Cochrane, S., Beyer, K., Clausen, M., Wjst, M. et al., Factors influencing the incidence and prevalence of food allergy. *Allergy* 2009, *64*, 1246–1255.

[4] Roses, J. B., Food allergen law and the Food Allergen Labeling and Consumer Protection Act of 2004: falling short of true protection for food allergy sufferers. *Food Drug Law J.* 2011, *66*, 225–242.

[5] Metcalfe, D. D., Astwood, J. D., Townsend, R., Sampson, H. A. et al., Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Crit. Rev. Food Sci. Nutr.* 1996, *36S*, S165–S186.

[6] Nordlee, J. A., Taylor, S. L., Townsend, J. A., Thomas, L. A. et al., Identification of a Brazil-nut allergen in transgenic soybeans. *N. Engl. J. Med.* 1996, *334*, 688–692.

[7] Goodman, R. E., Vieths, S., Sampson, H. A., Hill, D. et al., Allergenicity assessment of genetically modified crops—what makes sense? *Nat. Biotechnol.* 2008, *26*, 73–81.

[8] Goodman, R. E., Biosafety: evaluation and regulation of genetically modified (GM) crops in the United States. *J. Huazhong Agric. Univ.* 2014, *33*, 85–114.

[9] Goodman, R. E., Performing IgE serum testing due to bioinformatics matches in the allergenicity assessment of GM crops. *Food Chem. Toxicol.* 2008, *46*, S24–S34.

[10] Aalberse, R. C., Structural biology of allergens. *J. Allergy Clin. Immunol.* 2000, *106*, 228–238.

[11] Fernandez-Rivas, M., Fruit and vegetable allergy. *Chem. Immunol. Allergy* 2015, *101*, 162–170.

[12] ILSI-IFBC, Supplemental issue on the safety of foods and food ingredients developed through the application of genetic modification. *Reg. Toxicol. Pharma.* 1990, *12*, 1–196.

[13] FDA., Department of Health and Human Services: Food and Drug Administration (FDA): Statement of Policy: Foods Derived from New Plant Varieties. Federal Register Vol. 57, No. 104. Rockville, FDA, 1992.

[14] CODEX Alimentarius Commission., Proposed draft annex on the assessment of possible allergenicity of the draft guideline for the conduct of food safety assessment of foods derived from recombinant DNA plants. Joint FAO/WHO Food Standard Program, CAC/GL 45–2003, Appendix IV; 2003, pp. 57–60.

[15] EFSA Panel on Genetically Modified Organisms (GMO), Draft scientific opinion on the assessment of allergenicity of GM plants and microorganisms and derived food and feed. *EFSA J.* 2010, *8*. 1700, Appendix 3.

[16] Astwood, J. D., Fuchs, R. L., Lavric, P. B., in: Metcalfe, D. D., Sampson, H. A., Simmon, R., *Food Allergy*, Chapter 4, 2nd edn. Blackwell Sciences Inc., Cambridge, MA, USA 1997, pp. 65–92

[17] Goodman, R. E., Hefle, S. L., Taylor, S. L., van Ree, R., Assessing genetically modified crops to minimize the risk of increased food allergy: a review. *Int. Arch. Allergy Immunol.* 2005, *137*, 153–166.

[18] Goodman, R. E., Practical and predictive bioinformatics methods for the identification of potentially cross-reactive protein matches. *Mol. Nutr. Food Res.* 2006, *50*, 655–660.

[19] van Ree, R., Vieths, S., Poulsen, L. K., Allergen-specific IgE testing in the diagnosis of food allergy and the event of a positive match in the bioinformatics search. *Mol. Nutr. Food Res.* 2006, *50*, 645–654.

[20] Barrio, A. M., Soeria-Atmadja, D., Nister, A., Gustafsson, M. G. et al., EVALLER: a web server for in silico assessment of potential protein allergenicity. *Nuc. Acids Res.* 2007, *35*, 694–700.

[21] Oezguen, N., Zhou, B., Negi, S. S., Ivanciuc, O. et al., Comprehensive 3D-modeling of allergenic proteins and amino acid composition of potential conformational IgE epitopes. *Mol. Immunol.* 2008, *45*, 3740–3747.

[22] Dimitrov, I., Bangov, I., Flower, D. R., Doytchinova, I. AllerTOP v 2.—a server for in silico prediction of allergens. *J. Mol. Model.* 2014, *20*, 2278–6 DOI 10.1007/s00894-014-2278-5.

[23] Prosperi, M. C., Belgrave, D., Buchan, I., Simpson, A. et al., Challenges in interpreting allergen microarrays in relation to clinical symptoms: a machine learning approach. *Pediatr. Allergy Immunol.* 2014, *25*, 71–79.

[24] Pascal, M., Grishina, G., Yang, A. C., Sanchez-Garcia, S. et al., Molecular diagnosis of shrimp allergy: efficiency of several allergens to predict clinical reactivity. *J. Allergy Clin. Immunol. Pract.* 2015, *3*, 521–529.

[25] CODEX Alimentarius Commission. *Foods derived from modern biotechnology*, 2nd ed. FAO and WHO of the United Nations. CAC/GL 55–2003, CAC/GL 45–2003, CAC/GL 46–2003 and CAC/GL 68–2008. FAO Viale delle Terme di Caracalla, Rome, Italy 2009.

[26] EFSA. European Food Safety Authority Draft Scientific opinion on the assessment of allergenicity of GM plants and microorganisms and derived food and feed 2009.

[27] Goodman, R. E., Tetteh, A. O., Suggested improvements for the allergenicity assessment of genetically modified plants used in foods. *Curr. Allergy Asthma Rep.* 2011, *11*, 317–324.

[28] Cressman, R. F., Ladics, G., Further evaluation of the utility of "sliding window" FASTA in predicting cross-reactivity with allergenic proteins. *Regul. Toxicol. Pharmacol.* 2008, *54*, S20–S25.

[29] Silvanovich, A., Bannon, G., McClain, S., The use of *E*-scores to determine the quality of protein alignments. *Regul. Toxicol. Pharmacol.* 2009, *54*, S26–S31.

[30] Hoff, M., Son, D. Y., Gubesch, M., Ahn, K. et al., Serum testing of genetically modified soybeans with special emphasis on potential allergenicity of the heterologous protein CP4 EPSPS. *Mol. Nutr. Food Res.* 2007, *51*, 946–955.

[31] Holzhauser, T., van Ree, R., Poulsen, L. K., Bannon, G. A., Analytical criteria for performance characteristics of IgE binding methods for evaluating safety of biotech food products. *Food Chem. Toxicol.* 2008, *46*, S15–S19.

[32] King, T. P., Hoffman, D., Lowenstein, H., Marsh, D. G. et al., Allergen nomenclature. *Allergy* 1995, *50*, 765–774.

[33] Hileman, R. E., Silvanovich, A., Goodman, R. E., Rice, E. A. et al., Bioinformatics methods for allergenicity assessment using a comprehensive allergen database. *Int. Arch. Allergy Immunol.* 2002, *128*, 280–291.

[34] Sampson, H. A., Food allergy—accurately identifying clinical reactivity. *Allergy* 2005, *60*, 19–24.

[35] Bock, S. A., Munoz-Furlong, A., Sampson, H. A., Further fatalities caused by anaphylaxis to foods (Letter to the editor). *J. Allergy Clin. Immunol.* 2007, *119*, 1016–1018.

[36] Palmer, G. W., Dibbern, D. A., Burks, A. W., Bannon, G. A. et al., Comparative potency of Ara h 1 and Ara h 2 in immunochemical and functional assays of allergenicity. *Clin. Immunol.* 2005, *115*, 302–312.

[37] Pascal, M., Munoz-Cano, R., Reina, Z., Palacin, A. et al., Lipid transfer protein syndrome: clinical patter, cofactor effect and profile of molecular sensitization to plant-foods and pollens. *Clin. Exp. Allergy* 2013, *42*, 1529–1539.

[38] Le, T.-M., van Hoffen, E., Kummeling, I., Potts, J. et al., Food allergy in the Netherlands: differences in clinical severity, causative foods, sensitization and DBPCFC between community and outpatients. *Clin. Transl Allergy* 2015, *5*, 8, DOI: 10.1186/s.

[39] Worm, M., Eckermann, O, Dolle, S., Aberer, W. et al., Triffers and treatment of anaphylaxis: an analysis of 4,000 cases from Germany, Austria and Switzerland. *Dtsch. Arztebl. Int.* 2014, *111*, 367–375.

[40] Wong, G. W., Mahesh, P. A., Ogorodova, L., Leung, T. F. et al., The EuroPrevall-INCO surveys on the prevalence of food allergies in children from China, India and Russia: the study of methodology, *Allergy* 2010, *65*, 385–390.

[41] Incorvaia, C., Mauro, M., Ridolo, E., Makri, E. et al., A pitfall to avoid when using an allergen microarray: the incidental detection of IgE to unexpected allergens. *J. Allergy Clin. Immunol Pract.* 2015, *3*, 879–882.

[42] Palmer, G. W., Dibbern, D. A. Jr., Burks, A. W., Bannon, G. A. et al., Comparative potency of Ara h 1 and Ara h 2 in immunochemical and functional assays of allergenicity. *Clin. Immunol.* 2005, *115*, 302, 312.

[43] Vickery, B. P., Scurlock, A. M., Jones, S. M., Burks, A. W., Mechanisms of immune tolerance relevant to food allergy. *J. Allergy Clin. Immunol.* 2011, *127*, 576–584.

[44] Rosenfeld, L., Shreffler, W., Bardina, L, Niggemann, B. et al., Walnut allergy in peanut-allergic patients: significance of sequential epitopes of walnut homologous to linear epitopes of Ara h 1, 2 and 3 in relation to clinical reactivity. *Int. Arch. Allergy Immunol.* 2012, *157*, 238–245.

[45] Pastorello, E. A., Ortolani, C., Baroglio, C., Pravettoni, V. et al., Complete amino acid sequence determination of the major allergen of peach (*Prunus persica*) Pru p 1. *Biol. Chem.* 1999, *380*, 1315–1320.

[46] Fernandez-Rivas, M., Benito, C., Gonzalez-Mancebo, E., de Durana, D. A., Allergies to fruits and vegetables. *Pediatr. Allergy Immunol.* 2008, *19*, 675–681.

[47] Pascal, M., Munoz-Cano, R., Reina, Z., Palacin, A. et al., Lipid transfer protein syndrome: clinical, cofactor effect and profile of molecular sensitization of plant-foods and pollens. *Clin. Exp. Allergy* 2012, *42*, 1529–1539.

[48] Asero, R., Lipid transfer protein cross-reactivity assessment in vivo and in vitro in the office: pros and cons. *J. Investig. Allergol. Clin. Immunol.* 2011, *21*, 129–136.

[49] Swoboda, I., Jilek, A., Ferreira, F., Engel, E. et al., Isoforms of Bet v 1, the major birch pollen allergen, analyzed by liquid chromatography, mass spectrometry, and cDNA cloning. *J. Biol. Chem.* 1995, *270*, 2607–2613.

[50] Erdmann, S. M., Sachs, B., Schmidt, A., Merk, H. F. et al., In vitro analysis of birch-pollen-associated food allergy by use of recombinant allergens in the basophil activation test. *Int. Arch. Allergy Immunol.* 2005, *136*, 230–238.

[51] Thomas, W. R., Hierarchy and molecular properties of house dust mite allergens. *Allergol. Intl.* 2015, *64*, 304–311.

[52] Goodman, R. E., Leach, J. N., Assessing the allergenicity of proteins introduced into genetically modified crops

using human IgE assays. *J. AOAC Int.* 2004, *87*, 1423–1432.

[53] Orovitg, A., Guardia, P., Barber, D., de la Torre, F. et al., Enhanced diagnosis of pollen allergy using specific immunoglobulin E determination to detect major allergens and pan allergens. *J. Investig. Allergol. Clin. Immunol.* 2011, *21*, 253–259.

[54] Ferreira, F., Wolf, M., Wallner, M., Molecular approach to allergy diagnosis and therapy. *Yonsei Med. J.* 2014, *55*, 839–852.

[55] Chapman, M. D., Wuenschmann, S., King, E., Pomes, A., Technological innovations for high-throughput approaches to in vitro allergy diagnosis. *Curr. Allergy Asthma Rep.* 2015, *15*, 36.

[56] Crameri, R., Immunoglobulin E-binding autoantigens: biochemical characterization and clinical relevance. *Clin. Exp. Allergy* 2012, *42*, 343–351.

[57] Dema, B., Pellefigues, C., Hasni, S., Gault, N. et al., Autoreactive IgE is prevalent in systematic lupus erythematosus and is associated with increased disease activity and nephritis. *PLoS One* 2014, *9*: e90424, doi: 10.1371/journal.pone.0090424.

[58] Gavala, M. L., Bashir, H., Gern, J. E., Virus/allergen interactions in asthma. *Curr. Allergy Asthma Rep.* 2013, *13*, 298–307.

[59] Agrawal, R., Wisniewski, J., Yu, M. D., Kennedy, J. L. et al., Infection with human rhinovirus 16 promotes enhanced IgE responsiveness in basophils of atopic asthmatics. *Clin. Exp. Allergy* 2014, *44*, 1266–1273.

[60] Zhang, L., Liu, J., Wang, E., Zeng, S. et al., Respiratory syncytial virus protects against the subsequent development of ovalbumin-induced allergic responses by inhibiting Th2-type γδ T cells. *J. Med. Virol.* 2013, *85*, 149–156.

[61] Rossi, G. A., Colin, A. A., Infantile respiratory syncytial virus and human rhinovirus infections: respective role in inception and persistence of wheezing. *Eur. Respir. J.* 2015, *45*, 774–789.

[62] di Somma, C., Fiore, L., Di Lonardo, A., Ridolfi, B. et al., Cross-reactivity between the major *Parietaria* allergen and rotavirus VP4 protein. *Allergy* 2003, *58*, 503–510.

[63] Kelso, J. M., Mootrey, G. T., Tsai, T. F., Anaphylaxis from yellow fever vaccine. *J. Allergy Clin. Immunol.* 1999, *103*, 698–701.

[64] Farouk, S. E., Dolo, A., Berczky, S., Kouniba, B. et al., Different antibody-and cytokine-mediated reponses to *Plasmodium falciparum* parasite in two sympatric ethnic tribes living in Mali. *Microbes Infect.* 2005, *7*, 110–117.

[65] Halstensen, T. S., Why are we not all allergic: basic mechanisms for tolerance development. *Environ. Toxicol. Pharmacol.* 1997, *4*, 25–31.

[66] Arlian, L. G., Morgan, M. S., Quirce, S., Maranon, F. et al., Characterization of allergens in *Anisakis simplex*. *Allergy* 2003, *58*, 1299–1303.

[67] Hofer, M. F., Harbeck, R. J., Schlievert, P. M., Leung, D. Y., Staphylococcal toxins augment specific IgE responses by atopic patients exposed to allergen. *J. Invest. Dermatol.* 1999, *112*, 171–176.

[68] Wehner, J., Neuber, K., Staphylococcus aureus enterotoxins induce histamine and leukotriene release in patients with atopic eczema. *Br. J. Dermatol.* 2001, *145*, 302–305.

[69] Becker, Y., Respiratory syncytial virus (RSV) evades the human adaptive immune system by skewing the Th1/Th2 cytokine balance toward increased levels of Th2 cytokines and IgE, markers of allergy-a review. *Virus Genes* 2006, *33*, 235–252.

[70] Sampson H. A., Update on food allergy. *J. Allergy Clin. Immunol.* 2004, *113*, 805–819.

[71] Sicherer, S. H., Sampson, H. A., Food allergy: epidemiology, pathogenesis, diagnosis and treatment. *J. Allergy Clin. Immunol.* 2014, *133*, 291–307.

[72] Commins, S. P., Satinover, S. M., Hosen, J., Mozena, J. et al., Delayed anaphylaxis, angioedema, or urticaria after consumption of red meat in patients with IgE antibodies specific for galactose-α-1,3-galactose. *J. Allergy Clin. Immunol.* 2009, *123*, 426–433.

[73] Sampson, H. A., Historical background, definitions and differential diagnosis. *Chem. Immunol. Allergy* 2015, *101*, 1–7.

[74] Santos, A. F., Du Toit, G., Douiri, A., Badulovic, S. et al., Distinct parameters of the basophil activation test reflect the severity and threshold of allergic reactions to peanut. *J. Allergy Clin. Immunol.* 2015, *135*, 179–186.

[75] Mari, A., Ooievaar-de Heer, P., Scala, E., Giani, M. et al., Evaluation by double-blind placebo-controlled oral challenge of the clinical relevance of IgE antibodies against plant glycans. *Allergy* 2008, *63*, 891–896.

[76] Westphal, S., Kolarich, D., Foetisch, K, Lauer, I. et al., Molecular characterization and allergenic activity of Lyc e 2 (β-fructofuranosidase), a glycosylated allergen of tomato. *Eur. J. Biochem.* 2003, *270*, 1327–1337.

[77] Bublin, M., Dennstedt, S., Buchegger, M., Antonietta Ciardiello, M. et al., The performance of a component-based allergen microarray for the diagnosis of kiwifruit allergy. *Clin. Exp. Allergy* 2011, *41*, 129–136.

[78] Siruguri, V., Bharatraj, D. K., Vankudavath, R. N., Rao Mendu, V. V. et al., Evaluation of Bar, Barnase and Barstar recombinant proteins expressed in genetically engineered *Brassica juncea* (Indian mustard) for potential risks of food allergy using bioinformatics and literature searches. *Food Chem. Toxicol.* 2015, *83*, 93–102.

[79] Wang, F., Robotham, J. M., Teuber, S. S., Sathe, S. K. et al., Ana o 2, a major cashew (*Anacardium occidentale* L.) nut allergen of the legumin family. *Int. Arch. Allergy Immunol.* 2003, *132*, 27–39.

[80] Moneo, I., Caballero, M. L., Gonzalez-Munoz, M., Rodriguez-Mahillo, A. I. et al., Isolation of a heat-resistant allergen from fish parasite *Anisakis simplex*. *Parasitol. Res.* 2005, *96*, 285–289.

[81] Rodriguez-Mahillo, A. I., Gonzalez-Munoz, M., Gomez-Aguado, F., Rodriguez-Perez, R. et al., Cloning and characterization of the *Anisakis simplex* allergen Ani s 4 as a cysteine-protease inhibitor. *Int. J. Parasitol.* 2007, *37*, 907–917.

[82] Blank, S., Seismann, H., Bockisch, B., Braren, I. et al., Identification, recombinant expression, and characterization of the 100 kDa high molecular weight hymenoptera venom allergens Api m 5 and Ves v 3. *J. Immunol.* 2010, *184*, 5403–5413.

[83] Chua, K. Y., Doyle, C. R., Simpson, R. J., Turner, K. J. et al., Isolation of cDNA coding for the major mite allergen Der p II by IgE plaque immunoassay. *Int. Arch. Allergy Appl. Immunol.* 1990, *91*, 118–123.

[84] An, S., Chen, L., Long, C., Liu, X. et al., Dermatorphagoides farina allergens diversity identification by proteomics. *Mol. Cell Proteomics* 2013, *12*, 1818–1828.

[85] Postigo, I., Gutierrez-Rodriguez, A., Fernandez, J., Guisantes, J. A. et al., Diagnostic value of Alt a 1, fungal enolase and manganese-dependent superoxide dismutase in the component-resolved diagnosis of allergy to Pleosporaceae. *Clin. Exp. Allergy* 2011, *41*, 443–451.

[86] Gabriel, M. F., Postigo, I., Gutierrez-Rodriguez, A., Sunen, E. et al., Characterisation of *Alternaria alternata* manganese-dependent superoxide dismutase, a cross-reactive allergen homologue of Asp f 6. *Immunobiology* 2015, *220*, 851–858.

[87] Mattsson, L., Lundgren, T., Olsson, P., Sundberg, M. et al., Molecular and immunological characterization of Can f 4: a dog dander allergen cross-reactive with a 23 kDa odorant-binding protein in cow dander. *Clin. Exp. Allergy* 2010, *40*, 1276–1287.

[88] Niemi, M. H., Rytkonen-Nissinen, M., Janis, J., Virtanen, T. et al., Structural aspects of dog allergies: the crystal structure of a dog dander allergen Can f 4. *Mol. Immunol.* 2014, *61*, 7–15.

[89] Irifune, K., Nishida, T., Egawa, H., Nagatani, A., Pectin methylesterase inhibitor cDNA from kiwi fruit. *Plant Cell Rep.* 2004, *23*, 333–338.

[90] Kumari, D., Arora, N., Kasera, R., Sridhara, S. et al., Isolation and characterization of a 28 kDa major allergen from blackgram (*Phaseolus mungo*). *Immunobiology* 2012, *217*, 895–904.

Allergy EUROPEAN JOURNAL OF ALLERGY AND CLINICAL IMMUNOLOGY

Allergy

REVIEW ARTICLE

Update of the WHO/IUIS Allergen Nomenclature Database based on analysis of allergen sequences

C. Radauer[1,*], A. Nandy[2,*], F. Ferreira[3,*], R. E. Goodman[4,*], J. N. Larsen[5,*], J. Lidholm[6,*], A. Pomés[7,*], M. Raulf-Heimsoth[8,*], P. Rozynek[8], W. R. Thomas[9,*] & H. Breiteneder[1,*]

[1]Department of Pathophysiology and Allergy Research, Medical University of Vienna, Vienna, Austria; [2]Research and Development, Allergopharma GmbH & Co. KG, Reinbek, Germany; [3]Christian Doppler Laboratory for Allergy Diagnosis and Therapy, University of Salzburg, Salzburg, Austria; [4]Food Allergy Research & Resource Program, University of Nebraska Lincoln, Lincoln, NE, USA; [5]ALK A/S, Hørsholm, Denmark; [6]Thermo Fisher Scientific, Uppsala, Sweden; [7]Basic Research, Indoor Biotechnologies Inc., Charlottesville, VA, USA; [8]Institute of Prevention and Occupational Medicine of the German Social Accident Insurance, Ruhr University Bochum (IPA), Bochum, Germany; [9]TVW Telethon Institute for Child Health Research, University of Western Australia, West Perth, WA, Australia

To cite this article: Radauer C, Nandy A, Ferreira F, Goodman RE, Larsen JN, Lidholm J, Pomés A, Raulf Heimsoth M, Rozynek P, Thomas WR, Breiteneder H. Update of the WHO/IUIS Allergen Nomenclature Database based on analysis of allergen sequences. *Allergy* 2014; **69**: 413 419.

Keywords
allergen nomenclature; allergens; bioinformatics; database; protein families.

Correspondence
Heimo Breiteneder, Department of Pathophysiology and Allergy Research, Medical University of Vienna, Waehringer Guertel 18 20, 1090 Vienna, Austria.
Tel.: +43 1 40400 5102
Fax: +43 1 40400 5130
E mail: heimo.breiteneder@meduniwien.ac.at

*All of these authors belong to the WHO/IUIS Allergen Nomenclature Sub Committee.

Accepted for publication 12 November 2013

DOI:10.1111/all.12348

Edited by: Reto Crameri

Abstract

The IUIS Allergen Nomenclature Sub Committee, under the auspices of the World Health Organization and the International Union of Immunological Soci eties, maintains the systematic nomenclature of allergenic proteins and publishes a database of approved allergen names on its Web site, www.allergen.org. In this paper, we summarize updates of allergen names approved at the meetings of the committee in 2011 through 2013. These changes reflect recent progress in identifi cation, cloning, and sequencing of allergens. The goals of this update were to increase consistency in the classification of allergens, isoallergens, and variants and in the incorporation of the evolutionary classification of proteins into aller gen nomenclature, while keeping changes of established names to a minimum in the interest of continuity. Allergens for which names have been updated include respiratory allergens from birch and ragweed pollen, midge larvae, and horse dander; food allergens from peanut, cow's milk, and tomato; and cereal grain allergens. The IUIS Allergen Nomenclature Sub Committee encourages researchers to use these updated allergen names in future publications.

The official nomenclature of allergenic proteins is based on the Linnaean binominal nomenclature identifying genus and species of all organisms and was first published in 1986 (1) and revised in 1994 (2 6). The allergen nomenclature is maintained by the IUIS Allergen Nomenclature Sub Committee under the auspices of the World Health Organization (WHO) and the International Union of Immunological Societies (IUIS). The committee maintains the database of approved allergen names (www.allergen.org), which has developed from a plain text list to a fully functional, searchable database. In order to maintain a consistent allergen nomenclature that complies with the guidelines established by the subcommittee, researchers are required to submit newly described allergens to the Allergen Nomenclature Sub Committee before submitting their manuscript to a journal for consideration for publication. Submissions are kept confidential by the subcommittee, and no specific information other than the name of the new allergen will be disclosed on the Web site before publication. The submission form is available at www.allergen.org.

Allergen names are composed of an abbreviation of the scientific name of its source (genus: 3 4 letters; species: 1 2 letters) and an Arabic numeral, for example Der p 1 for the first allergen to be described from the house dust mite *Dermatophagoides pteronyssinus*. Originally, new allergens were assigned consecutive numbers. During the past decades, the increase in sequence data together with advances in bioinformatics made it possible to classify allergens into protein families whose members are evolutionary related, have

similar sequences and structures, and are, in some cases, also cross reactive (7 9). Hence, homologous allergens within a taxonomic order or family are now assigned corresponding numbers whenever possible, in order to reflect evolutionary relationships between allergens from different sources. For instance, in the rose family (Rosaceae), the numbers 1 5 are assigned to Bet v 1 related proteins (e.g., Mal d 1), thaumatin like proteins (Mal d 2), nonspecific lipid transfer proteins (Pru p 3), profilins (Mal d 4), and isoflavone reductases (Pyr c 5), respectively. These families of homologous allergens are frequently referred to as groups, although these designations are not part of the official allergen nomenclature. Examples are the group 1 mite allergens (e.g., Der p 1, Blo t 1, Eur m 1), which are papain like cysteine proteases, and the group 1 grass pollen allergens (e.g., Phl p 1, Lol p1, Cyn d 1), which are β expansins. Nevertheless, established numbers are usually not changed to avoid inconsistencies with allergen names used in previous publications or in allergen based products used in clinical practice. Bet v 1 related allergens from legumes, for instance, are named Ara h 8 in peanut, Gly m 4 in soybean, and Vig r 1 in mung bean.

Different closely related molecular species of an allergen are named by four digits following the period after the main allergen number. The first two digits designate isoallergens, which are defined as allergens from a single species with similar molecular masses, similar biochemical functions, and sequence identities >67%. The third and fourth digits distinguish different variants of an isoallergen, which are defined as proteins with more than 90% sequence identity. Variants with different nucleotide sequences encoding identical amino acid sequences do not receive individual designations. Both the 67% and the 90% identity thresholds represent arbitrary limits and serve merely as guidelines. Appropriate allergen designations are assigned on a case by case basis.

Table 1 Updated nomenclature of Bet v 1 isoallergens and variants. Boldface: updated allergen designations

Previous name	New name	UniProt	Other names	Comment
Bet v 1.0101	Bet v 1.0101	P15494	Bet v 1a	
Bet v 1.0102	**Bet v 1.0101**	P15494	Bet v 1 clone 224	Identical to Bet v 1.0101
Bet v 1.0103	**Bet v 1.0101**	P15494	Bet v 1 clone 2230	Identical to Bet v 1.0101
Bet v 1.0201	Bet v 1.0201	P45431	Bet v 1b	
Bet v 1.0301	**Bet v 1.0202**	P43176	Bet v 1c	
Bet v 1.0401	**Bet v 1.0102**	P43177	Bet v 1d	
Bet v 1.0402	**Bet v 1.0102**	P43177	Bet v 1h	Identical to 'Bet v 1.0401'
Bet v 1.0501	**Bet v 1.0103**	P43178	Bet v 1e	
Bet v 1.0601	**Bet v 1.0104**	P43179	Bet v 1f	
Bet v 1.0602	**Bet v 1.0104**	P43179	Bet v 1i	Identical to 'Bet v 1.0601'
Bet v 1.0701	**Bet v 1.0105**	P43180	Bet v 1g	
Bet v 1.0801	**Bet v 1.0106**	P43183	Bet v 1j	
Bet v 1.0901	**Bet v 1.0203**	P43184	Bet v 1k	
Bet v 1.1001	**Bet v 1.0107**	P43185	Bet v 1l	
Bet v 1.1101	(Deleted)	Q39417	Bet v 1 Sc1	Pathogen induced expression
Bet v 1.1201	(Deleted)	Q39420	Bet v 1 Sc2	Pathogen induced expression
Bet v 1.1301	(Deleted)	Q39415	Bet v 1 Sc3	Pathogen induced expression
Bet v 1.1401	**Bet v 1.0204**	P43186	Bet v 1m	
Bet v 1.1402	**Bet v 1.0204**	P43186	Bet v 1n	Identical to 'Bet v 1.1401'
Bet v 1.1501	(Deleted)	Q42499	BVGC11	Genomic sequence
Bet v 1.1502	(Deleted)	Q42499	BVGC31	Genomic sequence
Bet v 1.1601	(Deleted)	Q39425	BVGC181	Genomic sequence
Bet v 1.1701	(Deleted)	Q39426	BVGC21	Genomic sequence
Bet v 1.1801	(Deleted)	Q39427	BVGC25	Genomic sequence
Bet v 1.1901	(Deleted)	Q39428	BVGC34	Genomic sequence
Bet v 1.2001	(Deleted)	Q39429	BVGC45	Genomic sequence
Bet v 1.2101	(Deleted)	Q39430	BVGC63	Genomic sequence
Bet v 1.2201	(Deleted)	Q39431	BVGC681	Genomic sequence
Bet v 1.2301	(Deleted)	O23754	BVGC70	Genomic sequence
Bet v 1.2401	**Bet v 1.0108**	Q96365	Bet v 1 clone 167	
Bet v 1.2501	**Bet v 1.0109**	Q96366	Bet v 1 clone 184	
Bet v 1.2601	**Bet v 1.0110**	Q96367	Bet v 1 clone 2225	
Bet v 1.2701	**Bet v 1.0111**	Q96368	Bet v 1 clone 2226	
Bet v 1.2801	**Bet v 1.0112**	P15494 variant F63L	Bet v 1 clone 2227	
Bet v 1.2901	**Bet v 1.0113**	Q96370	Bet v 1 clone 2229	
Bet v 1.3001	**Bet v 1.0114**	Q96371	Bet v 1 clone 2301	

Researchers are encouraged to use full isoallergen and variant designations in order to unambiguously identify the allergens they work with. The importance of correct isoallergen/variant designations is highlighted by examples of highly different IgE binding and T cell activating properties of closely related isoallergens of Bet v 1 from birch pollen (10) and Der p 2 from house dust mite (11).

Updates of allergen designations

The 1994 revision of the allergen nomenclature represented the first introduction of bioinformatics into the allergen nomenclature, and sequence information became mandatory for the inclusion of new allergens. During the last three decades, bioinformatics has developed at an ever increasing speed and a very large amount of sequence data related to allergens has been generated. Many allergens recorded in the IUIS allergen database were originally submitted with partial sequences or even without associated sequence data. In most cases, full sequences have later become available, which in some cases has led to inconsistencies concerning the numbers assigned to allergen names. Therefore, the IUIS Allergen Nomenclature Sub Committee aimed to screen the database for such entries and to correct them based on sequences and data from the literature.

The database was manually searched for entries with conspicuous features such as missing sequence data, biochemical names similar to those of other allergens from the same source, or inconsistent allergen numbers compared with homologous allergens from the same taxonomic family. Allergen sequences were analyzed by pairwise and multiple sequence alignments. We found several types of incorrect allergen designations in the database and changed them according to the guidelines described above. A list of updated allergen designations is shown in Tables 1–4. These changes were approved by the IUIS Allergen Nomenclature Sub Committee at its meetings in the years 2011–2013.

Table 2 Updated nomenclature of pectate lyase allergens from short ragweed (*Ambrosia artemisiifolia*) pollen. Boldface: updated allergen designations

Previous name	New name	UniProt
Amb a 1.0101	Amb a 1.0101	P27759
Amb a 1.0201	Amb a 1.0201	P27760
Amb a 1.0202	Amb a 1.0202	E1XUL3
Amb a 1.0301	Amb a 1.0301	P27761
Amb a 1.0302	Amb a 1.0302	P27761 (variant L48Y)
Amb a 1.0303	Amb a 1.0303	P27761 (variant H392R)
Amb a 1.0304	Amb a 1.0304	E1XUL4
Amb a 1.0305	Amb a 1.0305	E1XUL5
Amb a 1.0401	Amb a 1.0401	P28744
Amb a 1.0402	Amb a 1.0402	E1XUL9
Amb a 2.0101	**Amb a 1.0501**	P27762
Amb a 2.0102	**Amb a 1.0502**	E1XUM1

Updated nomenclature of Bet v 1 isoallergens and variants

The major birch pollen allergen, Bet v 1, comprises a large number of isoallergens and variants, of which 36 were recorded in the IUIS allergen database (Table 1). A review of these entries revealed several problems: First, 13 entries referred to gene or cDNA sequences whose expression in pollen had not been demonstrated (12, 13). The sequences of Bet v 1.15–Bet v 1.23 were obtained from genomic DNA without determining their expression profiles (12). Bet v 1.11–Bet v 1.13 were identified in pathogen infected cell cultures and leaves, while no expression in pollen was shown (13). Hence, the entries Bet v 1.11–Bet v 1.13 and Bet v 1.15–Bet v 1.23 were deleted from the database.

Within the remaining 23 entries, four sets of isoallergens had identical protein sequences but different nucleotide sequences (Bet v 1.0101/02/03, Bet v 1.0401/02, Bet v 1.0601/02, Bet v 1.1401/02). In these cases, only the first

Table 3 Updated nomenclature of *Chironomus thummi thummi* hemoglobin allergens. Boldface: updated allergen designations

Previous name	New name	UniProt	Other names
Chi t 1.0101	Chi t 1.0101	P02229	Hemoglobin component III
Chi t 1.0201	Chi t 1.0201	P02230	Hemoglobin component IV
Chi t 2.0101	Chi t 2.0101	P02221	Hemoglobin component I
Chi t 2.0102	Chi t 2.0102	P02221 (variant A113T)	Hemoglobin component IA
Chi t 3.0101	Chi t 3.0101	P02222	Hemoglobin component II beta
Chi t 4.0101	Chi t 4.0101	P02231	Hemoglobin component IIIA
Chi t 5.0101	**Chi t 3.0201**	P02224	Hemoglobin component VI
Chi t 6.0101	**Chi t 3.0301**	P02226	Hemoglobin component VIIA
Chi t 6.0201	**Chi t 3.0401**	P02223	Hemoglobin component IX
Chi t 7*	**Chi t 3.0501**	P12548	Hemoglobin component VIIB 3
Chi t 7*	**Chi t 3.0601**	P84296	Hemoglobin component VIIB 4
Chi t 7*	**Chi t 3.0701**	P84298	Hemoglobin component VIIB 5/9
Chi t 7*	**Chi t 3.0702**	P12549	Hemoglobin component VIIB 6
Chi t 7*	**Chi t 3.0801**	P12550	Hemoglobin component VIIB 7
Chi t 8.0101	**Chi t 3.0901**	P02227	Hemoglobin component VIII
Chi t 9.0101	Chi t 9.0101	P02228	Hemoglobin component X

*The Uniprot entry P02225, previously listed in the IUIS allergen database for Chi t 7, was demerged into 7 entries, 5 from *C. thummi thummi* and 2 from *C. thummi piger*.

Table 4 Other updated allergen designations

Source	Previous name	New name	Uniprot acc. no.	Biochemical name
High sequence identity to another allergen from the same source				
Arachis hypogaea (peanut)	Ara h 4.0101	Ara h 3.0201	Q9SQH7	11S globulin; legumin; glycinin
High sequence identity to a homologous allergen from another source from the same taxonomic family				
Secale cereale (rye)	Sec c 1.0101	Sec c 38.0101	Q9S8H2	Dimeric α amylase/trypsin inhibitor
Hordeum vulgare (rye)	Hor v 21.0101	Hor v 20.0101	P80198	γ Hordein
Duplicate entries				
Hordeum vulgare (barley)	Hor v 1.0101	Hor v 15.0101	P16968	Monomeric α amylase inhibitor BMAI 1
Equus caballus	Equ c 5.0101	Equ c 4.0101	P82615	Latherin
Different proteins merged into a single allergen name				
Bos domesticus (cattle)	Bos d 8	Bos d 8		Whole casein fraction
		Bos d 9.0101	P02662	αS1 Casein
		Bos d 10.0101	P02663	αS2 Casein
		Bos d 11.0101	P02666	β Casein
		Bos d 12.0101	P02668	κ Casein
Update of botanical nomenclature				
Solanum lycopersicum (previously *Lycopersicon esculentum*; tomato)	Lyc e 1.0101	Sola l 1.0101	Q93YG7	Profilin
	Lyc e 2.0101	Sola l 2.0101	Q547Q0	β Fructofuranosidase
	Lyc e 2.0201	Sola l 2.0201	Q8RVW4	β Fructofuranosidase
	Lyc e 3.0101	Sola l 3.0101	P93224	Nonspecific lipid transfer protein
	Lyc e 4.0101	Sola l 4.0101	O49881	Pathogenesis related protein PR 10

named isoallergen designation was kept in the database and the multiple nucleotide sequence accession numbers coding for identical amino acid sequences were associated with these records. The remaining 18 sequences were grouped into two isoallergens: Bet v 1.01 with 14 variants and Bet v 1.02 with four variants. Different variants of each isoallergen showed 91–99% sequence identity, whereas identity between Bet v 1.01 and Bet v 1.02 sequences was 84–89% (Fig. 1A).

Different allergen numbers assigned to closely related allergens from the same source

The first two identified allergens from short ragweed (*Ambrosia artemisiifolia*) pollen were originally termed antigen E and antigen K (14, 15) and later renamed to Amb a 1 and Amb a 2 with the establishment of the IUIS nomenclature system (1). Both allergens belong to the pectate lyase family and show considerable IgE cross-reactivity (16). A sequence alignment yielded 61–70% sequence identity between the ten Amb a 1 variants (corresponding to four isoallergens) and the two Amb a 2 variants (Fig. 1B). Hence, Amb a 2.0101 and Amb a 2.0102 were renamed to Amb a 1.0501 and Amb a 1.0502, as two variants of a fifth Amb a 1 isoallergen (Table 2).

Freeze-dried larvae of the nonbiting midge *Chironomus thummi thummi* used as fish feed can elicit respiratory allergic reactions to its hemoglobin, which is composed of several distinct proteins originally named CTT I to CTT IX (17) and later designated *Chi t* I (18), following the original allergen nomenclature guidelines (1). As a consequence of the heterogeneous composition of *C. thummi thummi* hemoglobin, the Allergen Nomenclature Sub-Committee assigned a separate allergen name to each hemoglobin component in 2003, creating the allergen designations Chi t 1–Chi t 9. A multiple sequence alignment of these allergens showed sequence identities ranging from 28% to 99%. Particularly, isoallergens of Chi t 5–Chi t 8 showed 51–63% identity to Chi t 3.0101 (Fig. 1C). Despite having sequence identities below the 67% threshold, their identical molecular weights and biochemical functions justified the consolidation of these isoallergens into the single allergen name Chi t 3 (Table 3). The names of Chi t 1, Chi t 2, Chi t 4, and Chi t 9 remained unchanged. Hence, the number of *C. thummi thummi* hemoglobin allergens was reduced from 9 to 5.

Ara h 3 and Ara h 4 from peanut (*Arachis hypogaea*) are major allergens belonging to the 11S seed storage globulin family. Both were submitted to the Allergen Nomenclature Sub-Committee almost simultaneously with partial sequences and different molecular masses (19, 20). Subsequent molecular cloning revealed that their full-length sequences shared 91% identity, by far exceeding the 67% identity threshold for naming isoallergens. Consequently, Ara h 4.0101 was renamed to Ara h 3.0201 (Table 4).

A

	Bet v 1.0101	Bet v 1.0114	Bet v 1.0112	Bet v 1.0109	Bet v 1.0113	Bet v 1.0102	Bet v 1.0107	Bet v 1.0105	Bet v 1.0108	Bet v 1.0110	Bet v 1.0111	Bet v 1.0103	Bet v 1.0104	Bet v 1.0106	Bet v 1.0202	Bet v 1.0203	Bet v 1.0201	Bet v 1.0204
Bet v 1.0101	100	99	99	99	99	96	94	95	98	98	97	95	96	96	88	88	89	89
Bet v 1.0114	99	100	99	99	98	95	94	94	97	97	96	94	95	96	87	88	88	89
Bet v 1.0112	99	99	100	99	98	95	94	94	97	97	98	95	95	96	88	88	89	89
Bet v 1.0109	99	99	99	100	99	96	95	96	98	97	96	96	96	97	88	88	89	89
Bet v 1.0113	99	98	98	99	100	96	95	95	97	96	96	94	95	96	88	88	89	89
Bet v 1.0102	96	95	95	96	96	100	98	98	98	97	96	92	94	93	86	86	86	88
Bet v 1.0107	94	94	94	95	95	98	100	96	97	96	95	92	93	93	84	85	86	86
Bet v 1.0105	95	94	94	96	95	98	96	100	98	96	96	91	93	93	88	88	88	88
Bet v 1.0108	98	97	97	98	97	98	97	98	100	99	98	94	94	95	86	86	87	88
Bet v 1.0110	98	97	97	97	96	97	96	96	99	100	98	93	93	94	85	86	86	87
Bet v 1.0111	97	96	98	96	96	96	95	96	98	98	100	93	93	93	85	86	86	87
Bet v 1.0103	95	94	95	96	94	92	92	91	94	93	93	100	98	99	86	87	88	88
Bet v 1.0104	96	95	95	96	95	94	93	93	94	93	93	98	100	99	86	87	87	88
Bet v 1.0106	96	96	96	97	96	93	93	93	95	94	93	99	99	100	86	87	88	88
Bet v 1.0202	88	87	88	88	88	86	84	88	86	85	85	86	86	86	100	99	99	98
Bet v 1.0203	88	88	88	88	88	86	85	88	86	86	86	87	87	87	99	100	99	99
Bet v 1.0201	89	88	89	89	89	86	86	88	87	86	86	88	87	88	99	99	100	98
Bet v 1.0204	89	89	89	89	89	88	86	88	88	87	87	88	88	88	98	99	98	100

B

	Amb a 1.0101	Amb a 1.0201	Amb a 1.0202	Amb a 1.0301	Amb a 1.0302	Amb a 1.0303	Amb a 1.0304	Amb a 1.0305	Amb a 1.0401	Amb a 1.0402	Amb a 1.0501	Amb a 1.0502
Amb a 1.0101	100	75	75	75	76	75	76	76	82	82	66	66
Amb a 1.0201	75	100	100	87	87	86	87	87	73	73	69	69
Amb a 1.0202	75	100	100	87	87	87	87	87	73	73	69	69
Amb a 1.0301	75	87	87	100	100	100	99	99	72	72	70	70
Amb a 1.0302	76	87	87	100	100	99	99	99	72	72	70	70
Amb a 1.0303	75	86	87	100	99	100	100	99	72	72	70	70
Amb a 1.0304	76	87	87	99	99	100	100	100	72	72	70	70
Amb a 1.0305	76	87	87	99	99	99	100	100	72	72	70	70
Amb a 1.0401	82	73	73	72	72	72	72	72	100	99	62	61
Amb a 1.0402	82	73	73	72	72	72	72	72	99	100	61	61
Amb a 1.0501	66	69	69	70	70	70	70	70	62	61	100	99
Amb a 1.0502	66	69	69	70	70	70	70	70	61	61	99	100

C

	Chi t 3.0301	Chi t 3.0401	Chi t 3.0101	Chi t 3.0701	Chi t 3.0702	Chi t 3.0501	Chi t 3.0601	Chi t 3.0801	Chi t 3.0901	Chi t 3.0201	Chi t 1.0101	Chi t 1.0201	Chi t 2.0101	Chi t 2.0102	Chi t 9.0101	Chi t 4.0101
Chi t 3.0301	100	68	63	62	62	63	63	61	59	54	49	49	46	46	45	33
Chi t 3.0401	68	100	63	55	55	59	61	57	58	51	46	44	45	45	46	31
Chi t 3.0101	63	63	100	55	55	56	57	57	57	51	49	44	45	45	49	29
Chi t 3.0701	62	55	55	100	99	94	88	79	52	55	42	43	43	43	48	32
Chi t 3.0702	62	55	55	99	100	94	88	79	52	56	42	43	43	43	48	32
Chi t 3.0501	63	59	56	94	94	100	91	81	55	58	42	43	42	42	48	30
Chi t 3.0601	63	61	57	88	88	91	100	81	56	57	43	43	42	42	50	32
Chi t 3.0801	61	57	57	79	79	81	81	100	59	58	45	42	46	46	49	33
Chi t 3.0901	59	58	57	52	52	55	56	59	100	54	45	46	52	52	48	32
Chi t 3.0201	54	51	51	55	56	58	57	58	54	100	51	49	47	47	46	34
Chi t 1.0101	49	46	49	42	42	42	43	45	45	51	100	85	47	47	43	33
Chi t 1.0201	49	44	44	43	43	43	43	42	46	49	85	100	45	45	38	30
Chi t 2.0101	46	45	45	43	43	42	42	46	52	47	47	45	100	99	41	28
Chi t 2.0102	46	45	45	43	43	42	42	46	52	47	47	45	99	100	41	28
Chi t 9.0101	45	46	49	48	48	48	50	49	48	46	43	38	41	41	100	33
Chi t 4.0101	33	31	29	32	32	30	32	33	32	34	33	30	28	28	33	100

Figure 1 Percentage amino acid sequence identity matrices of Bet v 1 isoallergens and variants (A), Amb a 1 isoallergens and variants (B), and *C. thummi thummi* hemoglobin allergens (C). The values were calculated based on multiple sequence alignments of the amino acid sequences. Updated allergen names according to Tables 1 4 are used. Shadings in A and B indicate >90% identity. Values >50% (light gray) and >95% (dark gray) are shaded in C.

These three examples from ragweed, nonbiting midges, and peanut illustrate that the 67% identity threshold serves only as a reference guideline, whereas the nature of the allergen (biochemical function, molecular structure, molecular weight), in addition to the degree of identity with homologous allergens, takes equal priority for the assignment of allergen names.

Inconsistent allergen numbers regarding protein family memberships

Sec c 1 from rye (*Secale cereale*), a major baker's asthma allergen, is a dimeric bifunctional inhibitor of proteases and α amylases (21). However, within the grass family (Poaceae), the allergen number 1 is reserved for β expansin pollen allergens. The only other dimeric α amylase/protease inhibitor in the IUIS allergen database is Tri a 28 from wheat (*Triticum aestivum*). The sequences of the N terminal fragment of Sec c 1 and the C terminal fragment of Tri a 28 overlap by only 16 residues with 38% identity, whereas Tri a 28 is 100% identical to a different dimeric amylase inhibitor from rye (UniProt: C3VWW4). Hence, Sec c 1 showed no close relationship to any other cereal allergens and was renamed to Sec c 38, receiving the next available number within the Poaceae family (Table 4).

Hor v 21 (γ hordein) from barley (*Hordeum vulgare*) is a member of the γ prolamin subfamily, which comprises also γ secalins from rye and γ gliadins from wheat. The only other γ prolamin in the allergen database is Sec c 20 with two isoallergens: Sec c 20.0101 (γ 70 secalin) and Sec c 20.0201 (γ 35 secalin). As the allergen number 21 is already reserved for α/β prolamins such as Tri a 21, for the α/β gliadin from wheat, Hor v 21 was renamed to Hor v 20 (Table 4).

Duplicate database entries

Previously, the monomeric α amylase/protease inhibitor BMAI 1 from barley had been assigned the designations Hor v 1 and Hor v 15. As BMAI 1 is homologous to Tri a 15 from wheat (44% sequence identity), the designation Hor v 15 was maintained and Hor v 1 was deleted (Table 4).

The horse (*Equus caballus*) dander allergens Equ c 4 and Equ c 5 were originally submitted with partial sequence data and molecular masses of 18.7 and 16.7 kDa (22). After the full sequence of horse latherin became available, the partial sequences of both Equ c 4 and Equ c 5 matched this protein. The original description of Equ c 4 and Equ c 5 most likely referred to the glycosylated and nonglycosylated forms of latherin. Thus, the entry Equ c 5 was deleted (Table 4).

Updated nomenclature of casein components

The allergen Bos d 8, casein from cow's (*Bos domesticus*) milk, refers to a mixture of several dissimilar proteins. The casein fraction of milk proteins contains components that belong to two unrelated protein families, one family comprising αS1 , αS2 , and β caseins, while κ caseins constitute the other family (23). Even within the α/β casein family, sequence identities are below 15%. Hence, the entry Bos d 8 was demerged into four separate allergens: Bos d 9.0101 (αS1 casein), Bos d 10.0101 (αS2 casein), Bos d 11.0101 (β casein), and Bos d 12.0101 (κ casein; Table 4). The name Bos d 8, which is widely established and has been used in numerous publications and names of commercial diagnostic tests, was kept and designates the whole casein fraction. This example illustrates that the subcommittee takes into consideration not only taxonomic and other scientific aspects but also practical aspects of continuity and public acceptance in its management of the allergen nomenclature.

Adjustment of allergen names to reflect updated taxonomy

In addition to the changes described above, based on sequence similarities and protein family memberships, the IUIS Allergen Nomenclature Sub Committee changed the designations of tomato allergens from Lyc e 1 Lyc e 4 to Sola l 1 Sola l 4 in order to reflect the establishment of *Solanum lycopersicum* instead of *Lycopersicon esculentum* as the official scientific name of the tomato (24) (Table 4).

Concluding remarks

The WHO/IUIS Allergen Nomenclature Database is continuously updated and supplemented not only with newly submitted allergens, but also with data of already published allergens. The IUIS Allergen Nomenclature Sub Committee encourages users to notify the committee of missing or inconsistent records in the database and thereby aid in providing a reliable and up to date resource of unambiguous allergen names and isoallergen/isoform sequence information for the scientific community. Updates and error reports may be sent to the subcommittee either by using the submission form to be downloaded from www.allergen.org or by directly contacting one of the committee members, whose contact details are published at www.allergen.org.

Acknowledgments

The authors thank all members of the WHO/IUIS Allergen Nomenclature Sub Committee for fruitful discussions and John Wise (Food Allergy Research & Resource Program, University of Nebraska Lincoln, Lincoln, NE, USA) for designing and maintaining the WHO/IUIS Allergen Nomenclature Database.

Author contributions

C. Radauer drafted the manuscript and performed the greater part of the sequence analyses. A. Nandy contributed to sequence analysis of Bet v 1 and Amb a 1 isoallergens. M. Raulf Heimsoth and P. Rozynek performed sequence analysis of *Chironomus* allergens. H. Breiteneder, F. Ferreira, R. Goodman, J. N. Larsen, J. Lidholm, A. Pomés, and W. R. Thomas contributed to discussions of several groups of allergens. All authors critically read and approved the manuscript.

Conflicts of interest

All authors have no conflicts of interest to declare.

References

1. Marsh DG, Goodfriend L, King TP, Lowenstein H, Platts-Mills TA. Allergen nomenclature. *Bull World Health Organ* 1986;**64**: 767 774.
2. King TP, Hoffman D, Lowenstein H, Marsh DG, Platts-Mills TA, Thomas W. Allergen nomenclature. *Int Arch Allergy Immunol* 1994;**105**:224 233.
3. King TP, Hoffman D, Lowenstein H, Marsh DG, Platts-Mills TA, Thomas W. Allergen nomenclature. *Bull World Health Organ* 1994;**72**:797 806.
4. King TP, Hoffman D, Lowenstein H, Marsh DG, Platts-Mills TA, Thomas W. Allergen nomenclature. *Clin Exp Allergy* 1995;**25**: 27 37.
5. King TP, Hoffman D, Lowenstein H, Marsh DG, Platts-Mills TA, Thomas W. Allergen nomenclature. *Allergy* 1995;**50**:765 774.
6. King TP, Hoffman D, Lowenstein H, Marsh DG, Platts-Mills TA, Thomas W. Allergen Nomenclature. *J Allergy Clin Immunol* 1995;**96**:5 14.
7. Chapman MD, Pomes A, Breiteneder H, Ferreira F. Nomenclature and structural biology of allergens. *J Allergy Clin Immunol* 2007;**119**:414 420.
8. Ferreira F, Hawranek T, Gruber P, Wopfner N, Mari A. Allergic cross-reactivity: from gene to the clinic. *Allergy* 2004;**59**:243 267.
9. Radauer C, Bublin M, Wagner S, Mari A, Breiteneder H. Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *J Allergy Clin Immunol* 2008;**121**:847 852.
10. Ferreira F, Hirtenlehner K, Jilek A, Godnik-Cvar J, Breiteneder H, Grimm R et al. Dissection of immunoglobulin E and T lymphocyte reactivity of isoforms of the major birch pollen allergen Bet v 1: potential use of hypoallergenic isoforms for immunotherapy. *J Exp Med* 1996;**183**:599 609.
11. Hales BJ, Hazell LA, Smith W, Thomas WR. Genetic variation of Der p 2 allergens: effects on T cell responses and immunoglobulin E binding. *Clin Exp Allergy* 2002;**32**:1461 1467.
12. Hoffmann-Sommergruber K, Vanek-Krebitz M, Radauer C, Wen J, Ferreira F, Scheiner O et al. Genomic characterization of members of the Bet v 1 family: genes coding for allergens and pathogenesis-related proteins share intron positions. *Gene* 1997;**197**: 91 100.
13. Swoboda I, Scheiner O, Heberle-Bors E, Vicente O. cDNA cloning and characterization of 3 genes in the Bet v 1 gene family that encode pathogenesis-related proteins. *Plant, Cell Environ* 1995;**18**:865 874.
14. King TP, Norman PS, Lichtenstein LM. Isolation and characterization of allergens from ragweed pollen. IV. *Biochemistry* 1967;**6**:1992 2000.
15. Adolphson C, Goodfriend L, Gleich GJ. Reactivity of ragweed allergens with IgE antibodies. Analyses by leukocyte histamine release and the radioallergosorbent test and determination of cross-reactivity. *J Allergy Clin Immunol* 1978;**62**:197 210.
16. Wopfner N, Gadermaier G, Egger M, Asero R, Ebner C, Jahn-Schmid B et al. The spectrum of allergens in ragweed and mugwort pollen. *Int Arch Allergy Immunol* 2005;**138**:337 346.
17. Baur X, Dewair M, Fruhmann G, Aschauer H, Pfletschinger J, Braunitzer G. Hypersensitivity to chironomids (non-biting midges): localization of the antigenic determinants within certain polypeptide sequences of hemoglobins (erythrocruorins) of Chironomus thummi thummi (Diptera). *J Allergy Clin Immunol* 1982;**69**:66 76.
18. Mazur G, Baur X, Liebers V. Hypersensitivity to hemoglobins of the Diptera family Chironomidae: structural and functional studies of their immunogenic/allergenic sites. *Monogr Allergy* 1990;**28**:121 137.
19. Kleber-Janke T, Crameri R, Appenzeller U, Schlaak M, Becker WM. Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology. *Int Arch Allergy Immunol* 1999; **119**:265 274.
20. Rabjohn P, Helm EM, Stanley JS, West CM, Sampson HA, Burks AW et al. Molecular cloning and epitope analysis of the peanut allergen Ara h 3. *J Clin Invest* 1999;**103**:535 542.
21. Garcia-Casado G, Armentia A, Sanchez-Monge R, Sanchez LM, Lopez-Otin C, Salcedo G. A major baker's asthma allergen from rye flour is considerably more active than its barley counterpart. *FEBS Lett* 1995;**364**:36 40.
22. Goubran Botros H, Poncet P, Rabillon J, Fontaine T, Laval JM, David B. Biochemical characterization and surfactant properties of horse allergens. *Eur J Biochem* 2001;**268**:3126 3136.
23. Ginger MR, Grigor MR. Comparative aspects of milk caseins. *Comp Biochem Physiol B Biochem Mol Biol* 1999;**124**:133 145.
24. Spooner DM, Anderson GJ, Jansen RK. Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes, and pepinos (Solanaceae). *Am J Botany* 1993;**80**:676 688.

Flexible sequence similarity searching with the **FASTA3** program package

William R. Pearson

Department of Biochemistry,
University of Virginia,
Charlottesville, VA 22908

August 28, 1998

Phone: 804-924-2818; FAX: 804-924-5069; wrp@virginia.edu

1. INTRODUCTION

Since the publication of the first rapid method for comparing biological sequences 15 years ago (1), DNA and protein sequence comparison have become routine steps in biochemical characterization, from newly cloned proteins to entire genomes. As the DNA and protein sequence databases become more complete, a sequence similarity search is more likely to reveal a database sequence with statistically significant similarity, and thus inferred homology, to a query sequence. Indeed, even in the archaebacterium *M. jannaschii*, more than 40% of the open reading frames could be assigned a function based on significant sequence similarity to a protein of known function (2).

This chapter provides a "hands on" overview of the programs in the FASTA package. Rather than discuss in depth the theory and practice of protein and DNA sequence comparison, I focus on more practical questions, such as: "Which FASTA program should I use?", "What threshold should I use for statistical significance?", "Which databases should I search?", "When should I use FASTA and when should I use BLAST?", and "When should I change the scoring matrix and gap penalties?" For an excellent review of similarity searching with BLAST and FASTA and of local similarity statistics, see ref. 3. For more specific information on how to use the FASTA programs to identify distantly related sequences, see refs. 4 and 5. A detailed explanation of the statistical estimates in the `fasta3` package is provided in ref. 6.

2. SIMILARITY SEARCHING WITH THE FASTA3 PROGRAMS

The FASTA program package has evolved significantly since its introduction ten years ago (7). The original package offered four programs: `fasta`, `tfasta`, `lfasta`, and `rdf` (`rdf` was introduced with the first `fastp` program in 1985; ref. 8). Today, programs are available for rigorous Smith- Waterman searches (`ssearch3`) and for searches with mixed peptide sequences (`fastf3` and `tfastf3`); the programs for translated DNA:protein sequence comparison have been improved substantially with the introduction of `fastx3`, `fasty3`, `tfastx3`, and `tfasty3`, and the program for estimating statistical significance from shuffled-sequence similarity scores (`prss3`) produces accurate statistical estimates. The FASTA3 programs for database searching are summarized in Table 1; the programs for evaluating statistical significance are shown in Table 2.

Table 1: Comparison programs in the FASTA3 package

`fasta3`	Compare a protein sequence to a protein sequence database or a DNA sequence to a DNA sequence database using the FASTA algorithm (4, 7). Search speed and selectivity are controlled with the ktup (wordsize) parameter. For protein comparisons, *ktup= 2* by default; *ktup=1* is more sensitive but slower. For DNA comparisons, *ktup= 6* by default; *ktup=3* or *ktup= 4* provides higher sensitivity; *ktup=1* should be used for oligonucleotides (DNA query lengths <20).
`ssearch3`	Compare a protein sequence to a protein sequence database or a DNA sequence to a DNA sequence database using the Smith-Waterman (22) algorithm. `ssearch3` is about 10-times slower than `FASTA3`, but is more

	sensitive for full-length protein sequence comparison.
`fastx3`/ `fasty3`	Compare a DNA sequence to a protein sequence database, by comparing the translated DNA sequence in three frames and allowing gaps and frameshifts. `fastx3` uses a simpler, faster algorithm for alignments that allows frameshifts only between codons; `fasty3` is slower but produces better alignments with poor quality sequences because frameshifts are allowed within codons.
`tfastx3`/ `tfasty3`	Compare a protein sequence to a DNA sequence database, calculating similarities with frameshifts to the forward and reverse orientations.
`tfasta3`	Compare a protein sequence to a DNA sequence database, calculating similarities (without frameshifts) to the 3 forward and three reverse reading frames. `tfastx3` and `tfasty3` are preferred because they calculate similarity over frameshifts.
`fastf3`	Compare a mixed peptide sequence to a protein sequence database. A mixture of peptides, typically obtained by Edman degradation after cyanogen bromide cleavage without further separation, is compared with protein sequences in a database to identify those sequences that are most likely to produce the peptide mixture.
`tfastf3`	Compare a mixed peptide sequence to a translated DNA sequence database.

Table 2: Statistics programs in the `FASTA3` package

`prss3`	Evaluate the significance of a protein or DNA sequence similarity score by comparing two sequences and calculating optimal similarity scores, and then repeatedly shuffling the second sequence, and calculating optimal similarity scores using the Smith-Waterman algorithm. The characteristic parameters of the extreme value distribution are estimated from the shuffled sequence scores and used to calculate the statistical significance of the unshuffled sequence similarity score.
`sc_to_e`	Calculate the statistical significance of a similarity score from the raw score, the length of the sequence, the statistical parameters estimated from a search, and the size of the database.
`randseq`	Produce a random sequence with the same length and amino acid composition as a query sequence. Random sequences are useful in evaluating the accuracy of statistical estimates. In general in a database search, the highest scoring match to a random query sequence should have an expectation value E() ~1.

In addition, several programs in the FASTA2 package are not yet included with the FASTA3 programs (Table 3). As this chapter is written (summer, 1998), `lalign` is the most important program in the FASTA2 package that is not in the `fasta3` package. `lalign` (and the related graphical programs `plalign` and `flalign`) can produce multiple local alignments from the same pair of protein sequences, while `fasta3` and `fasta` produce only one alignment. Multiple local alignments can highlight domains with proteins; i.e. a protein may contain several domains

that share strong similarity with a library sequence. When multiple similar domains are present, `fasta3` only shows the most similar alignment; `lalign` is required to detect the alternative alignments.

In general, programs in the FASTA3 package are preferred over the older FASTA2 programs if FASTA3 has the function you need. Programs in the FASTA3 package have more robust statistical estimates and error handling, a larger variety of scoring matrices (`fasta3` has MDM10, MDM20, PAM120, and BLOSUM80 in addition to PAM250, BLOSUM50, and BLOSUM62 in `fasta2`), and a broader array of comparison functions (`fasty3`, `fastf3`, `tfasty3`, and `tfastf3`).

Table 3: Programs available only with FASTA2

`lalign/ plalign/ flalign`	Find multiple local alignments between two protein or DNA sequences using the `sim` implementation (23) of the Waterman-Eggert (24) algorithm. `lalign` shows traditional alignments; `plalign` produces graphics, while `flalign` produces graphics commands for the GCG figure program. This program performs successive full Smith-Waterman alignments, and is best used for protein alignments. For DNA, try `lfasta` (below).
`lfasta/ plfasta/ flfasta`	Find multiple local alignments between two protein or DNA sequences using the `fasta` algorithm. `lalign` uses the heuristic `fasta` algorithm with a local band-alignment. `lalign` is preferred for protein alignment, but `lfasta` is much faster for very long DNA sequences. `plfasta` and `flfasta` produce graphical output.
`prdf`	Like `prss3`, but uses the `fasta` algorithm instead of Smith-Waterman. `prss3` is preferred.
`align`	Global sequence alignment between two protein or DNA sequences using linear space (25).
`aacomp`	Reports amino acid composition and molecular weight of a protein sequence.
`grease/ tgrease`	Calculates the hydropathy plot of a protein sequence using the Kyte-Doolittle method (26). `tgrease` produces tektronix graphics.

2.1 Which Program Should I Use?

Many investigators who use the `fasta` program for protein and DNA database searches are unfamiliar with other programs in the package, or are unclear as to when they should be used. Table 4 suggests some strategies for using the programs in the FASTA3 package.

The suggestions in Table 4 are based on two rules-of-thumb: (1) use the program that is designed for your problem; and (2) whenever possible, search protein sequence databases before DNA sequence databases. Protein sequence comparison routinely reveals homologous sequences that diverged 2-3 billion years ago; it is difficult for DNA sequence comparison to “look-back”

more than 200-500 million years. Thus, protein sequence comparison, or translated DNA sequence comparison, allows one to identify homologs that diverged 5-10-times farther back in evolutionary time (Table 5).

Table 4: Which Program When?

Problem	Program	Explanation	Alternative
Identify unknown protein	(1) `fasta3`	General protein comparison. Use *ktup=2* (the unknown default) for speed; *ktup=1* for a more sensitive search. Search first against the smallest library likely to contain a homolog (i.e. SwissProt rather than Genpept).	blastp/
	(2) `ssearch3`	10-50-fold slower than `fasta3`, but provides maximum sensitivity. No advantage for DNA comparisons.	`fasta3/ blastp`
	(3) `tfastx3/ tfasty3`	If a homolog cannot be found in the protein databases, check the DNA databases with `tfastx3` or `tfasty3`. `tfasty3` provides more accurate alignments, but is about 33% slower.	`tblastn/ tfasta`[a]
Identify structural DNA sequence	`fasta3`	If the DNA sequence encodes a protein, use protein sequence comparison first, then try translated protein sequence comparison (`fastx3`/`fasty3`). For repeated DNA sequences or structural RNAs, search first with *ktup=6* (the default), then *ktup=3*. Search with *ktup< 3* only for very short sequences (PCR primers).	`blastn`
Identify EST sequence	`fastx3/ fasty3`	Protein sequence comparison is far more sensitive than DNA comparison, so check first to see if the EST encodes a product homologous to a known protein. Current version searches forward strand only, so use `fastx3 -i` as well.	`fasta3/ blastx/ tblastx`
Identify new orthologs	`tfastx3/ tfasty3`	If possible, search EST sequences from the same species. Use low/close MDM20 scoring matrices to detect close relationships and avoid distant relationships. Confirm statistical significance	`tblastn/ tblastx`
Confirm statistical significance	`prss3`	Use 500-2000 shuffles, and remember to normalize the statistical significance to the size of the database originally searched (typically 10,000 - 100,000 sequences).	
Confirm statistical estimates	`randseq`	Use to generate random sequences; then search using `fasta3` (or `blastp` or `ssearch3`) and look for E() ~1.0.	

[a]No longer recommended.

In addition, low complexity regions are relatively easily removed from protein sequence databases and recognized in protein sequence alignments, but they are much more difficult to recognize in DNA sequence alignments. These regions can produce statistically significant similarity scores for non-homologous sequences because of their unusual amino-acid

composition. Thus, when seeking to identify a newly sequenced EST (Expressed Sequence Tag) sequence, you should first use `fastx3` or `fasty3` to search a comprehensive protein database like SwissProt or PIR, then search a larger but more redundant database like the BLAST/NCBI `nr` or OWL (9) "non-redundant" protein databases, or Genpept, and, only after these searches have failed to turn up statistically significant matches should you look for DNA sequence matches.

Table 5: DNA vs. protein sequence comparison

The best scores are:		DNA E(188,018)	`tfastx3` E(187,524)	prot. E(331,956)
DMGST	D.melanogaster GST1-1	1.3e-164	4.1e-109	1.0e-109
MDGST1	M.domestica GST-1 gene	2e-77	3.0e-95	1.9e-76
LUCGLTR	Lucilia cuprina GST	1.5e-72	5.2e-91	3.3e-73
MDGST2A	M.domesticus GST-2 mRNA	9.3e-53	1.4e-77	1.6e-62
MDNF1	M.domestica nf1 gene. 10	4.6e-51	2.8e-77	2.2e-62
MDNF6	M.domestica nf6 gene. 10	2.8e-51	4.2e-77	3.1e-62
MDNF7	M.domestica nf7 gene. 10	6.1e-47	9.2e-77	6.7e-62
AGGST15	A.gambiae GST mRNA	3.1e-58	4.2e-76	4.3e-61
CVU87958	Culicoides GST	1.8e-41	4.0e-73	3.6e-58
AGG3GST11	A.gambiae GST1-1 mRNA	1.5e-46	2.8e-55	1.1e-43
BMO6502	Bombyx mori GST mRNA	1.1e-23	8.8e-50	5.7e-40
AGSUGST12	A.gambiae GST1-1 gene	2.3e-16	4.5e-46	5.1e-37
MOTGLUSTRA	Manduca sexta GST	5.7e-07	2.5e-30	8.0e-25
RLGSTARGN	R.legominosarum *gstA* and *gstR*	0.0029	3.2e-13	1.4e-10
HUMGSTT2A	H. sapiens GSTT2	0.32	3.3e-10	2.0e-09
HSGSTT1	H.sapiens GSTT1 mRNA	7.2	8.4e-13	3.6e-10
ECAE000319	E. coli hypothet. prot.	—	4.7e-10	1.1e-09
MYMDCMA	Methylophilus dichlorometh. DH	—	1.1e-09	6.9e-07
BCU19883	Burkholderia maleylacetate red.	—	1.2e-09	1.1e-08
NFU43126	Naegleria fowleri GST	—	3.2e-07	0.0056
SP505GST	Sphingomonas paucim	—	1.8e-06	0.0002
EN1838	H. sapiens maleylacetoacetate iso.	—	2.1e-06	5.9e-06
HSU86529	Human GSTZ1	—	3.0e-06	8.0e-06
SYCCPNC	Synechocystis GST	—	1.2e-05	9.5e-06
HSEF1GMR	H.sapiens EF1g mRNA	—	9.0e-05	0.00065

The primate, other mammal, invertebrate, and bacterial sections of Genbank were searched using a *Drososphila* glutathione transferase cDNA (`DMGST`) and protein (`gtt1_drome`) sequence using `fasta3` (DNA, *ktup=4*), `tfastx3`, and `fasta3` (protein, *ktup=2*). Expectation values for selected high scoring sequences are shown. DNA comparisons with "—" had expectation values E()>100 .With this query, DNA sequence comparison detects homologs only in other insects, while protein and translated DNA comparison finds statistically significant similarity with homologs from humans and bacteria.

2.2 FASTA vs. BLAST

The BLAST family of sequence comparison programs (10, 11) offers many of the same search capabilities as the FASTA programs (Table 6). In general, the BLAST programs are faster, but the FASTA programs can provide more accurate alignments. For most protein sequence database searching, the current `blastp2.0` (gapped blast, ref. 11) will identify an unknown protein as effectively as `fasta3` and even the more rigorous `ssearch3`. `fasta3` and `ssearch3` use different scoring matrices (BLOSUM50) and gap penalties (-12 for the first residue in a gap, -2 for each additional residue) from `blastp2.0` (BLOSUM62, -12 for the first residue in a gap, -1 for each additional residue). The previous `blastp1.4` produced very poor sequence alignments (because of the restriction on gaps); but the current `blastp2.0` version produces protein alignments that are very similar to those obtained with a rigorous Smith-Waterman search.

Table 6: Comparison of BLAST2 and FASTA3 Programs

Program		
BLAST	FASTA	Function
`blastp`	`fasta3`	General protein sequence similarity searches. blastp is faster and can show alignments between several domains in the same sequence. `fasta3` displays a Smith-Waterman final alignment and produces more accurate statistical estimates in some cases.
`blastn`	`fasta3`	DNA sequence comparison. `blastn` is highly optimized for speed; it uses a fixed word size (11 nucleotides) and scoring matrix that are inappropriate for some problems (e.g. searching for PCR primer matches). `blastn` searches with both strands of a DNA sequence. `fasta3` does not; two searches (`fasta3` and `fasta3 -i`) are required.[a]
`blastx`	`fastx3`/ `fasty3`	Compare a translated DNA to a protein sequence database. While `blastx` does six independent searches (one for each of the six frames), `fastx3` and `fasty3` effectively does a single forward (or backward) search, which allows frameshifts in computing the similarity score and alignments. As a result, `fastx3` and `fasty3` are more sensitive and can produce much better alignments than blastx when the DNA sequence has frameshift errors. blastx searches in the forward and reverse frames; `fastx3`/`fasty3` searches only in the forward or the reverse (`fasty3 –i`) frame.
`tblastn`	`tfastx3`/ `tfasty3`/ `tfasta`	Compare a protein sequence to a DNA sequence database, translating in the three forward and reverse frames. Again, `tfastx3` and `tfasty3` provide more accurate alignments than `tblastn` or `tfasta` when the DNA sequences have frameshift errors.
	`tblastx`	Compare a DNA query sequence to a DNA library, translating both sequences in all six frames and scoring using a protein substitution matrix (BLOSUM62). `fasta3` with *ktup=6* (the default) provides a similar function, but does not use a protein scoring matrix.

[a]The GCG implementation of `fasta` searches with both strands.

For translated DNA-protein comparison and DNA database searches, the FASTA programs are much better than their BLAST counterparts. Although the gapped `blastp2.0` performs very well in protein comparisons, `blastx` performs the three forward-frame searches separately, while `fastx3` and `fasty3` calculate a single alignment that allows frameshifts. Treating the all three forward reading frames as a single sequence makes it much easier to produce high quality alignments that extend across the length of the matched protein sequence and allows similarity from the different reading frames to be combined in a natural way to improve sensitivity. For example, a `blastx` search with a class-mu mouse glutathione transferase cDNA sequence with insertion and deletion errors at 5% of the positions detected only other class-mu glutathione transferases, while a search with the same sequence using `fasty3` detected more class-mu protein sequences with $10^{-20} < E() < 10^{-17}$ and an additional 8 more distantly related class-pi glutathione transferase sequences ($10^{-5} < E() < 0.01$).

The FASTA programs also provide additional flexibility for DNA sequence searches. Searches can be done with any "wordsize" (*ktup*) from 1-6; small *ktup*'s are particularly appropriate for searches with short sequences, such as PCR primers. In addition the FASTA programs can use a variety of scoring matrices, including matrices with very high mismatch penalties that can be used to identify long identities in sequences.

3. INTERPRETING FASTA STATISTICS

When rapid sequence comparison programs were first introduced in 1983 (1), it became possible to find similar DNA and protein sequences by searching sequence databases, but there was no formal basis for deciding whether a weak similarity was likely to be biologically significant. A Monte-Carlo shuffling method for evaluating similarity scores (`rdf`) was provided with the FASTP program (8), but the recommended guidelines for significant similarity ($Z > 5$) were not based on the correct statistical model for local similarity scores and did not account for database size. A sequence with a score that is 10 standard deviations ($Z > 10$) above the mean is expected 0.015 times by chance in a search of a 10,000 entry database; the same score would be expected 0.11 times by chance in a search of SwissProt (70,000 entries), and thus would not be statistically significant, even at the 0.05 level.

Accurate statistical estimates were introduced into similarity searching with the `blastp` program (10), based on the recognition that local similarity scores can be described accurately by the extreme value distribution (12, 13). The Monte-Carlo shuffling program introduced with `fastp` now uses the extreme value distribution to calculate the probability of an alignment score, and the library searching programs in the FASTA2 and FASTA3 packages provide a value that can be used to infer homology from statistically significant similarity the expectation (E()) value (6).

The E() value is the first number that you should look at when deciding whether to analyze further a high-ranking sequence alignment. Investigators often wonder what E() value they should use. This is discussed in detail in the next section, but in most cases, and E() value between 0.001 and 0.01 can be used to infer homology reliably, but lower (more conservative) values are required when hundreds or thousands of searches are performed (as when characterizing all the genes in a bacterial genome).

The E()-value calculated by the `fasta3` programs and `blast` programs is a statistical measure of the likelihood that the observed similarity score could have occurred by chance. Like any statistical measure, its usefulness depends on: (1) whether the assumptions of the underlying statistical model are correct, and (2) the kinds of errors that one is willing to accept when using the measure to draw a conclusion. For similarity searching, we infer homology (common ancestry) from "statistically significant" similarity. However, the threshold for "statistical significance" will vary, depending on whether we are more concerned about occasionally mis-identifying a non-homolog (labeling a sequence as related when it is not, a false positive or type I error) or missing a likely homolog (labeling a sequence as non-homologous when a high-scoring homolog has been found, a false-negative or type II error).

3.1 What threshold should I use to infer sequence homology?

For most molecular biologists, the greatest concern in similarity searching is a false-positive error; we don't want to send a letter to Nature identifying a yeast homolog of `p53_human` when no evolutionary relationship exists.[1] While incorrect assertion of homology was relatively common before accurate similarity statistics became available, it is rare today. (Unfortunately however, once the "observation" has been published, it is difficult to remove from the literature.) The E()-value or expectation calculated by `fasta3` *et al*. is the number of times you would expect to see a score equal or greater by chance in a search of the database. In other words, E() < 0.01 says that you expect to see a score that high (or higher) once by chance in 100 searches; E() < 0.001 says once in 1000 searches, etc. E() ~1 says that you expect to see a score that high, simply by chance, every time you do a search.

Older versions of the `blast` programs used a related statistic, the *p()*-value, to characterize the significance of a similarity score. The E()-value reported by the fasta programs ranges from *0..D*, where *D* is the number of entries in the database, while the `blast` *p()*-value ranges from 0..1. The probability (*p()*-value) of an E()-value can be found with the Poisson formula: $p(E)=1-e^{-E}$. For values of E() < 0.1, *p()* ~ E(), thus *p*(E = 0.1) = 0.1; *p*(E = 1.0) = 0.63; *p*(E = 5.0) = 0.99.

While a sensible E()-value threshold (0.001 - 0.01) can ensure that researchers avoid "false positive" errors, little can be done to avoid "false negatives," i.e. labeling a sequence as unrelated to anything in the database when in fact a homolog is present. Most diverse protein families contain pairs of related sequences that do not share statistically significant sequence similarity. Fortunately, if those families are large (e.g. globins, serine proteases, glutathione transferases, G-protein coupled receptors), it is likely that newly discovered family members will share significant similarity with some known members of the family. As the sequence databases grow more complete and protein families expand, the rate of false negatives should decrease.

3.2 Choosing a database

The expectation value E(S > x) of a similarity score is calculated from the probability of the pair-wise similarity score *p*(S > x), which can be calculated using the extreme value distribution (12, 13), and the number of "tests" (i.e. sequence comparisons) that were performed to find the

[1]The gold-standard test for homology is structural similarity. If the candidate yeast homolog of P53 has a completely different three-dimensional structure, the hypothesis is wrong.

high-scoring sequence. Thus, E(S > x) = *p*(S > x)*D*, where *D* is the number of sequences in the database. (For DNA sequence comparison, *D* is not the number of sequences in the database but the length of the database in nucleotides divided by the length of the query sequence.)

Because E() increases linearly with the number of database entries, a similarity found in a search of a bacterial genome with 1,000-5,000 entries will be 50-250-fold more significant than an alignment with exactly the same score found in the OWL non-redundant protein database (ref. 9; 250,000 entries). Thus, when searching for very distant relationships, one should always use the smallest database that is likely to contain the homolog of interest. If the goal is to find the *E. coli* homolog of the *B. subtilis* DAHP synthase (`arog_bacsu`), one should search the *E. coli* proteome (which finds the E. coli *kdsA* homolog with E(4,283) < 0.00015) rather than SwissProt (`kdsa_ecoli` E(74,417) < 0.0017) or OWL (`kdsa_ecoli` E(260,784) < 0.0085). Here, the same alignment, with the same similarity score, is 50-fold less significant against the largest database than with the smallest.

Likewise, a search of SwissProt (~70,000 entries) will be 3-5 fold more sensitive than either OWL (261,000 sequences) or the BLAST `nr` protein database (332,000 sequences), simply because Swissprot is smaller. Thus, an efficient strategy for identifying protein homologs should: (1) search smaller databases first; then (2) re-search a smaller database (like SwissProt) with a more sensitive algorithm (`fasta3` with *ktup=1* or `ssearch3`), and then, if no significant matches are found, (3) search larger databases (OWL or `nr`).

While their size reduces search sensitivity, larger databases can be effective when they provide more diverse members of a protein family. For example, the most distant `p53_human` homolog in SwissProt is a flounder sequence. OWL contains about twice as many novel p53 homologs, including one from squid.

3.3 Thresholds for large-scale sequence analysis

Genome sequencing centers and other groups that do thousands of similarity searches each day must use more conservative thresholds of statistical significance to avoid false positive errors. A threshold of E() = 0.001, which is conservative for someone who does a few searches a day, should produce 10 scores below the threshold between non-homologous sequences by chance after 10,000 searches. Indeed, if you do 100 searches with random sequences against the PIR or Swissprot databases, one of those 100 sequences will find a "homolog" with E() < 0.01, ten will have E() < 0.1, etc. (6). Genome sequencing centers typically use thresholds of E()<10^{-6}, or even lower, when characterizing thousands of sequences.

However, using a more conservative threshold of statistical significance ensures that you will make more false negative (type II) errors when looking at distant relationships. For example, in a comparison of 2608 human proteins from SwissProt against the E. coli proteome (4289 sequences), 417 obtained E() < 0.02, 373 had E() < 0.01, 301 had E() < 0.001, 256 had E() < 0.0001. Of the 72 with 0.001 < E() < 0.01, we would expect that about 26 (0.01 2608) shared similarity this high by chance, while the other 45 are truly homologous. (Unfortunately, we cannot identify which 45 sequences are homologs without additional information.) In the human/E. coli search, 209 sequences had E()<10^{-6}; we would expect all of these matches are genuine homologies. However, using the conservative 10^{-6} threshold would misidentify as "unrelated" almost 200 probable homologs. Thus, estimates of the number of "novel" or

"unidentified" proteins in newly sequenced bacterial genomes are generally overestimates, since many of these "novel" proteins may share significant similarity when searched individually, but not when searched in a group of 2,000-4,000 sequences.

3.4 Statistical estimates–what can you trust?

If the statistical estimates are accurate, the guidelines in the previous section provide a reliable strategy for identifying related sequences based on sequence similarity. However, with biological sequences (as opposed to "fair" coins), the assumptions underlying the statistical model may not be met. When the assumptions fail, the highest scoring unrelated sequence may have an expectation value that is much too low (e.g. E()<10^{-3}) or much too high (E() > 100). If the E()-value is too low, unrelated sequences will be mistakenly labeled as related (false positives). If the E()-values are too high, it is likely that the E()-values of related sequences are too high as well, and related sequences will be missed (false negatives).

In general, inaccurate statistical estimates are caused by either (1) incorrect gap penalties or (2) low complexity regions (runs of simple amino acid composition, e.g. `ggqgppgdaggpg` from a *C. elegans* collagen or `ssggvtfsvss` from a *Drosophia* trypsin) in the query sequence (3, 14). In the first case, the statistical model has failed. The statistical theory behind the estimates for BLASTP, FASTA and Smith-Waterman (`ssearch3` scores assumes that the scores are "local," i.e. on average, non-identical amino acids will have similarity scores $s_{ij} < 0$. If the gap penalties are too low, then the alignment algorithm will choose to insert a gap, rather than to end the alignment, and the alignment will tend to become "global," aligning the sequences from end to end. The statistical properties of "global" alignment scores are different from those of "local" scores. "Local" scores follow the extreme-value distribution; the distribution of "global" alignment scores is not well understood.

The reliability of the sequence statistics can be confirmed quickly by looking at the histogram of observed and expected similarity scores that is displayed after a `fasta3` search,[2] and by checking the expectation (E()) value of the highest scoring unrelated sequence.[3] If there is good agreement between the observed and expected distribution of scores and the E() value of the highest scoring unrelated sequence is ~1, the statistical estimates should be accurate.

---

[2]These examples show results from running the `fasta3` and `ssearch3` programs, which are distributed from `ftp://ftp.virginia.edu/pub/fasta/`. The programs available from this site run on most UNIX platforms (Digital UNIX, IBM AIX, Linux, SGI Irix, and Sun Solaris) as well as Windows (Windows95 and NT) and Macintosh. The output shown here may differ slightly from the FASTA program distributed with the Genetics Computer Group, but similar information is available from all modern FASTA implementations.

[3]Although identifying the highest scoring unrelated sequence seems to presume knowledge of the protein family, additional searches with candidate unrelated sequences (E() ~1) can often separate low scoring related from high scoring unrelated sequences (5).

Figure 1: Histogram of `fasta3` similarity scores

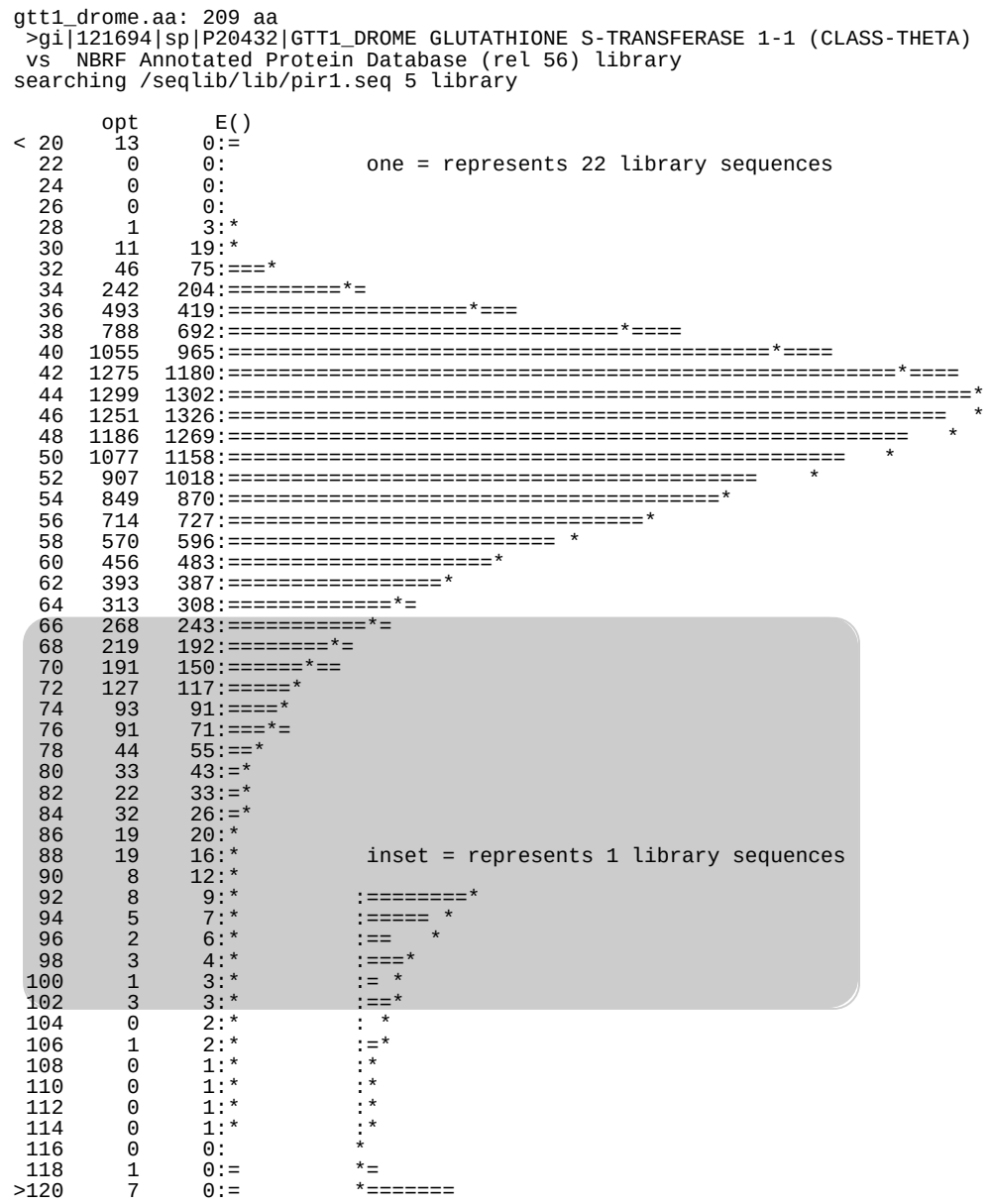


Fig. 1: Histogram of `fasta3` similarity scores—Results of a search of a *Drosophila* class-theta glutathione transferase (`gtt1_drome`) against the annotated PIR1 protein sequence database. The initial histogram output is shown. The shaded section indicates the region that is most likely to show discrepancies between observed and expected numbers of scores when the statistical model fails.

3.4.1 Low gap penalties cause inaccurate estimates

For most protein and DNA sequence searches, there is excellent agreement between the observed and expected distribution of scores (Fig. 1) and the E()-value of the highest scoring unrelated sequence is ~1.0 (Table 7; ref. 6). The FASTA programs provide a histogram summarizing the distribution of observed and expected scores after every search (Figs. 1-3). Fig. 1, reports that for this search, 788 sequences ("`opt`" column) in the database obtained scores of 38-39 (left-most column), while 692 sequences ("`E()`" column) are expected to have scores in that range for a database of 14,000 sequences. Agreement between observed ("`===`" graph) and expected ("`*`" in histogram) is especially important in the shaded area in Fig. 1. For many searches, it is also possible to confirm the accuracy of the estimates by looking for the highest scoring unrelated sequence in the list of high scoring sequences. In Table 7 the highest scoring unrelated sequences are `S30223` and `NOBY2`, with expectation values ~8. (Ideally, these scores would be a bit closer to 1; the highest scoring unrelated sequence in the same search with `ssearch3` has E() < 3.)

Table 7: FASTA search - high scoring sequences

Name	description	len	initn	opt	z-score	E()
XUFF11	glutathione transferase	209	1399	1399	1626.5	1.2e-84
XUZM32	glutathione transferase	222	133	173	210.9	8.6e-06
XUZM31	glutathione transferase	220	107	164	200.6	3.2e-05
XUZM1	glutathione transferase	213	123	144	177.7	0.00061
RGECSS	string. starv. prot. - E. coli	212	106	140	173.1	0.0011
XURTG	glutathione transferase	222	58	139	171.7	0.0013
XURT8C	glutathione transferase	222	39	115	144.0	0.046
XURTG4	glutathione transferase	218	40	93	118.7	1.2
A37378	glutathione transferase	210	40	82	106.2	5.8
S30223	elongation factor eEF-1g	227	34	80	103.5	8.3
NOBY2	*phosphopyruvate hydratase*	437	53	83	103.1	8.8
PWBYD	*H+-transporting ATP synthase*	212	53	79	102.7	9.2

High scoring sequences from searches of `gtt1_drome` against the annotated PIR1 database (27) with `fasta3` (*ktup=2*. High-scoring unrelated sequences are highlighted in *italics*.

Tables 8 and 9, and Fig. 2 show two examples of searches where the statistical model has failed. In the first case (Table 8), a DNA search was performed with gap penalties of -12 and -2, rather than the default -16, -4. While the histogram (not shown) shows good agreement between the observed and expected distribution of scores, the E()-value of the highest scoring unrelated sequence is 0.01. (That the high-scoring unrelated sequence does not contain a homolog was confirmed by scanning it with `tfasty3`). Moreover, the E()-values for homologous alignments increase by 10^7 (e.g. from 1.2×10^{-12} to 0.0008 for AC002520; Table 8) when the gap penalties are reduced from –16/-4 to –12/-2. DNA sequence searches with even lower gap penalties do show sizeable differences between the observed and expected distribution of scores, but the E()-value of the highest unrelated sequence is usually the most sensitive measure of the accuracy of the statistical estimates.

Figure 2: Poor statistics: low complexity regions

```
grou_drome.aa: 719 aa
 >GROU_DROME GROUCHO PROTEIN (ENHANCER OF SPLIT M9/10). - DROSOPHILA MELANOGAS
 vs  NBRF Annotated Protein Database (rel 56) library
searching /seqlib/lib/pir1.seq 5 library

       opt      E()
< 20    13     0:=
  22     0     0:           one = represents 28 library sequences
  24     1     0:=
  26     0     0:
  28     1     3:*
  30    10    20:*
  32    21    76:= *
  34   105   205:====   *
  36   272   422:==========     *
  38   540   697:====================    *
  40   937   972:==================================*
  42  1269  1188:=========================================*===
  44  1645  1311:================================================*============
  46  1666  1335:================================================*============
  48  1577  1278:============================================*===========
  50  1310  1166:=======================================*=====
  52  1056  1025:===================================*=
  54   851   876:============================*
  56   669   732:========================  *
  58   423   601:================     *
  60   419   487:===============  *
  62   255   390:==========   *
  64   196   310:=======    *
  66   181   245:======= *
  68   154   193:======*
  70    99   151:==== *
  72    74   118:=== *
  74    63    92:===*
  76    60    72:==*
  78    47    56:=*
  80    48    43:=*
  82    36    33:=*
  84    33    26:*=
  86    27    20:*
  88    21    16:*          inset = represents 2 library sequences
  90    18    12:*
  92    20     9:*         :====*=====
  94    20     7:*         :===*======
  96    17     6:*         :==*======
  98     7     4:*         :=*==
 100    10     3:*         :=*===
 102    11     3:*         :=*====
 104    10     2:*         :*====
 106    11     2:*         :*=====
 108     7     1:*         :*===
 110    10     1:*         :*====
 112     6     1:*         :*==
 114     4     1:*         :*=
 116    11     0:=         *======
 118    10     0:=         *=====
>120    70     0:===       *==================================
5446221 residues in 14321 sequences
 Expectation_n fit: rho(ln(x))= 8.0964+/-0.00108; mu= 4.7475+/- 0.061;
 mean_var=157.6967+/-31.622, 0's: 13 Z-trim: 96  B-trim: 33 in 1/62
 Kolmogorov-Smirnov  statistic: 0.0497 (N=29) at  52
```

Fig. 2: Poor statistics: low complexity regions—A fasta3 search (*ktup=2*) of the PIR1 database using grou_drome. The histogram of sequence similarity scores is shown. In this case, there are clear discrepancies between the observed and expected numbers of sequences with scores in the central part of the distribution and in the tails, and there is an excess of high scoring sequences. Table 9 shows that all of these excess high-scoring sequences are unrelated.

Table 8: FASTA search - low gap penalties

The best scores are:		(length)	initn	opt	z-sc	E(-12/-2)	E(-16/-4)
AC002520	Human Chr. 1p13	(11901)	1507	404	173.1	0.0008	1.2e-12
AC000031	Human Chr. 1p13.3	(39043)	1396	394	161.0	0.0011	6.5e-12
HSU47924	*Human chr. 12p13*	(78864)	235	352	138.3	0.01	2.0
AC000032	Human Chr. 1p13	(29867)	1354	345	141.6	0.018	6.6e-09
CACD42	*C.atys CD4 mRNA*	(1189)	69	307	146.1	0.26	—
HUMDXS455A	*Human cosmid*	(38409)	126	274	109.2	0.89	—
HSHS12ENH	*Homo sapiens DNA*	(3735)	151	278	126.1	1.1	0.038
HSV411C11	*Human DNA*	(5637)	165	276	122.5	1.1	—
HUMHSLA	*Human hormone-sens.*	(3255)	63	275	125.7	1.3	—
AF031078	*Human chr. X*	(78864)	188	264	100.2	1.4	0.078
AF035180	*Human chr. 4q35*	(4638)	67	271	121.7	1.5	0.08

High-scoring sequences from a fasta3 search (*ktup=6*) of the Primate division of Genbank 106 (~80,0000 sequences) using the reverse complement of a m*Gstm1* cDNA sequence (MUSGLUTA) using the default substitution matrix (+5/-4) and low (-12/-2) or default (-16/-4) gap penalties. Unrelated sequences are highlighted with *italics*. The low gap penalties improve the E()-value of the unrelated HSU47924 sequence to E() < 0.01 and reduce the significance of the homologous AC002520, AC000031, and AC000032 sequences by 10^7 .

3.4.2 Low E-values from low-complexity regions

Low E()-values between non-homologous sequences are usually caused by low complexity regions (3,14). The *Drosophila* "groucho" protein sequence (grou_drome) contains only 5 low complexity regions (83 of 719 residues as determined by seg, ref. 14), but as comparison of Fig. 2 and Fig. 3 shows, matches in these regions significantly distort the distribution of the high-scoring unrelated sequences. In contrast, a search with the 5 low-complexity regions masked (Fig. 3) shows the expected distribution of scores. Examination of the list of high-scoring sequences in the low- complexity search (Table 9) shows a large number of "significant" matches (0.00013 < E() < 0.02) to unrelated proteins with biased amino-acid compositions, while the highest scoring unrelated sequence in the "seg-ed" search has E() < 0.047 Perhaps surprisingly, the significance of the related GTP-binding regulatory protein similarity scores improve almost 1000-fold as well (Table 9).

Table 9: FASTA search – low complexity regions

Search with complete grou_drome:		length	initn	init1	opt	z-sc	E(14,212)
RGHUB1	GTP-binding reg. prot.	(340)	161	147	237	197.4	4.9e-05
RGHUB3	GTP-binding reg. prot.	(340)	163	152	233	194.2	7.4e-05
RGBOB2	GTP-binding reg. prot.	(326)	181	149	228	190.5	0.00012
PIHUB6	*salivary proline-rich prot*	(392)	142	142	229	190.1	0.00013
RGKWB	GTP-binding reg. prot.	(340)	159	154	222	185.4	0.00023
RGFFBH	GTP-binding reg. prot.	(340)	169	144	219	183.0	0.00031
PIHUSD	*proline-rich glycoprot.*	(310)	141	141	217	182.0	0.00035
PIRT3	*acidic proline-rich protein*	(206)	138	138	212	180.7	0.00042
WMBEW6	*capsid protein - herpes*	(635)	101	101	206	168.7	0.002
S23447	*annexin XI form B-bovine*	(505)	84	84	202	166.9	0.0024
PIHUPF	*salproline-rich glycoprot.*	(251)	147	147	193	164.3	0.0034
PIHUSC	*proline-rich phosphoprot.*	(166)	88	88	180	156.6	0.0092

Search with complete grou_drome:		length	initn	init1	opt	z-sc	E(14,212)
CGHU6C	*collagen alpha 1 (II)*	(1487)	104	104	197	156.0	0.0099
RGOOBE	GTP-binding reg. prot.	(341)	156	125	181	152.8	0.015
FOLJSP	*gag polyprotein - foamy vir*	(811)	121	121	187	151.9	0.017
CGBO1S	*collagen alpha 1 (I)-bovine*	(779	88	88	185	150.6	0.02
LUDO7	*annexin VII - slime mold*	(462)	88	88	179	149.2	0.024
CGHU2S	*collagen alpha 2 (I)*	(1366)	88	88	187	148.6	0.026
LUBO11	*annexin XI form A-bovine*	(503)	84	84	177	147.1	0.031
S09257	*Hox A4 - chicken*	(309)	116	116	172	146.2	0.035
OZZQMY	*circumsporozoite prot pre.*	(367)	146	146	172	145.1	0.04
Search with seg-ed grou_drome: (low complexity regions removed)							
RGHUB1	GTP-binding reg. prot.	(340)	161	147	237	247.5	8e-08
RGHUB3	GTP-binding reg. prot.	(340)	163	152	233	243.3	1.4e-07
RGHUB2	GTP-binding reg. prot.	(340)	181	149	228	238.1	2.7e-07
RGKWB	GTP-binding reg. prot.	(340)	159	154	222	231.9	5.9e-07
RGFFBH	GTP-binding reg. prot.	(340)	169	144	219	228.7	8.9e-07
RGOOBE	GTP-binding reg. prot.	(341)	156	125	181	189.1	0.00014
BVBYMS	*MSI1 protein - yeast*	(422)	116	74	139	143.9	0.047
ERHUAH	*coatomer complex alpha*	(1224)	109	109	134	131.7	0.23
I37062	*involucrin S - gorilla*	(495)	129	81	115	117.8	1.3

Unrelated sequences are highlighted in *italics.*

For protein-protein database searches, removal of low-complexity sequences is equally effective for either the query sequence or the protein database. However, it is more difficult to remove low-complexity regions from DNA query sequences, such as EST sequences. Unfortunately, high-scoring alignments between low-complexity protein sequences and out-of-frame DNA translations are common (15). A simple strategy for improving the sensitivity of translated DNA searches (fastx3, fasty3, or blastx) is to search against a "seg-ed" protein database (14).

Low-gap penalties and low-complexity regions produce unreliable statistical estimates because the underlying assumptions of the statistical model do not apply. Low gap penalties cause alignments to shift from local to global; extreme-value alignment statistics apply only to local alignments. Low-complexity regions violate implicit assumptions about higher-order structure in the "unrelated" sequences. With low-complexity sequences the matches are statistically significant but not biologically significant, because the statistical model assumed that each position of a random (unrelated sequence) is independent of all the others.

When the statistical model is valid—local alignments and truly "random" unrelated sequences—statistically significant similarity scores can be used to infer homology reliably. And one can usually check that the statistical model is correct by looking at the histogram of observed and expected similarity scores, and by checking the expectation value of the highest scoring unrelated sequence.

Figure 3: Accurate statistics with “seg-ed” query

```
grou_drome.seg: 719 aa
 >GROU_DROME GROUCHO PROTEIN (ENHANCER OF SPLIT M9/10). - DROSOPHILA MELANOGAS
 vs  NBRF Annotated Protein Database (rel 56) library
searching /seqlib/lib/pir1.seq 5 library

       opt      E()
< 20    48     0:==
  22    14     0:=          one = represents 24 library sequences
  24    21     0:=
  26    37     0:==
  28    39     3:*=
  30    65    20:*==
  32    95    76:===*
  34   175   206:========*
  36   348   424:===============  *
  38   591   700:========================    *
  40   891   977:====================================  *
  42  1141  1194:================================================ *
  44  1328  1317:=========================================================*=
  46  1373  1342:==========================================================*==
  48  1395  1285:============================================================*=====
  50  1227  1172:=============================================*===
  52  1107  1031:=========================================*====
  54   888   880:====================================*
  56   723   735:==============================*
  58   602   604:=========================*
  60   490   489:====================*
  62   357   392:=============== *
  64   284   312:============*
  66   246   246:==========*
  68   177   194:========*
  70   131   152:======*
  72   110   119:====*
  74    64    93:===*
  76    76    72:==*=
  78    53    56:==*
  80    41    43:=*
  82    44    33:=*
  84    22    26:=*
  86    26    20:*=
  88    17    16:*          inset = represents 1 library sequences
  90    11    12:*
  92    14     9:*         :========*=====
  94     5     7:*         :===== *
  96     7     6:*         :=====*=
  98    11     4:*         :===*=======
 100     2     3:*         :==*
 102     5     3:*         :==*==
 104     3     2:*         :=*=
 106     1     2:*         :=*
 108     1     1:*         :*
 110     0     1:*         :*
 112     1     1:*         :*
 114     0     1:*         :*
 116     0     0:          *
 118     1     0:=         *=
>120    13     0:=         *==============
5446221 residues in 14321 sequences
 Expectation_n fit: rho(ln(x))= 6.3481+/-0.00105; mu= 10.5411+/- 0.059;
 mean_var=92.0111+/-17.844, 0's: 13 Z-trim: 24  B-trim: 593 in 1/62
 Kolmogorov-Smirnov  statistic: 0.0129 (N=29) at  42
```

Fig. 3: Accurate statistics with “seg-ed” query—The search in Fig. 3 was performed using the grou_drome sequence with low-complexity sequences masked using the “seg” program (14). With low complexity sequences removed, the numbers of observed and expected similarity scores agree closely. Identical results are obtained when low-complexity regions are removed from the PIR1 database instead of grou_drome.

Table 10: FASTA3 general options

-a	show full sequences rather than only overlapping region (fastx/y3 and tfastx/y3 do not provide this feature)
-b #	number of best scores to show (must be < -E cutoff)
-d #	number of best alignments to show (must be < -E cutoff)
-E #	Expectation value limit for displaying scores and alignments. (By default, 10.0 for protein sequence comparisons; 5.0 for fastx/y3, and 2.0 for DNA sequence comparisons.)
-H	turn off histogram display
-I	(DNA only) reverse complement the query sequence; by default `fasta3`, `fastx3`, and `ssearch3` search only with the forward sequence. (tfastx/y3) compare against only the reverse complement of the library sequences.
-L	report long sequence description in alignments
-m 1-6,10	alignment display options (Table 14)
-n	force query to nucleotide sequence (default: autodetect)
-N #	read database in chunks of # residues. # should be > 2-times the query sequence length, as the chunks overlap by the length of the query. (default: 80,000 query-length)
-O file	send output to file
-q/-Q	quiet option; do not prompt for input
-r file	save all scores to statistics file
-S #	offset substitution matrix values
-s name	scoring matrix. BLOSUM50 is used by default for proteins, PAM120, PAM250, and BLOSUM62 can be specified by setting -s P120, P250, or BL62. Additional matrices include: BLOSUM80 (BL80), and MDM_10, MDM_20, MDM_40 (M10, M20, M40, 19). Alternatively, BLASTP1.4 for- mat scoring matrix files can be specified.
-w #	line width for similarity score and sequence alignment output
-W #	amount of sequence context around the alignment. Default is 30 residues (not used by fastx/y3, tfastx/y3).
-x "#,#"	offsets query and library sequence for numbering alignments
-z #	specify statistics calculation. Default is -z 1. Table 13.
-Z #	specify the size of the library to be used for statistical significance estimates.

4. FASTA3 PROGRAM OPTIONS

The behavior of the programs in the FASTA package can be modified with a variety of command line options; options are available to change the scoring matrix and gap penalties, use alternate statistical estimation methods, and change the format of the alignment output. Many of the options apply to all of the programs in the package (Table 10); other options are specific to `fasta3` or tfastx/y3 (Table 11). When using the FASTA programs distributed from the U. of Virginia, command line options must precede other program arguments. The standard invocation of a FASTA program is:

```
program -opt1 -opt2 arg2 -opt3 query_file library ktup-opt
```

specifically:

```
fasta3 -q -f -14 -w 75 -L -m 1 mgstm1.aa /slib/swissprot 1
```

In the latter case, the `fasta3` program is run in “quiet” (-q) mode with a penalty for the first residue in a gap of -14 (`-f -14` rather than the default -12), alignments are printed at 75 residues per line (`-w 75`), a long description of the library sequence is shown with the alignment (`-L`), and the alignment symbol highlights the differences rather than similarities (`-m 1`). Fig. 4

shows the difference between a conventional alignment (Fig. 4A) and one produced with the command line options shown above (B).

Table 11:

`fasta3`, `fastx/y3`, `tfastx/y3`, `tfasta3` options	
`-1`	sort by "init1" score
`-3`	(`tfasta3`, `tfastx3`, `tfasty3` only) use only forward frame translations
`-A`	force Smith-Waterman alignment for output. Smith-Waterman is the de fault for protein sequences, `fastx/y3`, and `tfastx/y3`, but not for `tfasta3` or DNA comparisons with `fasta3`.
`-c #`	threshold for band optimization
`-f #`	penalty for the first residue in a gap
`-g #`	penalty for additional residues in a gap
`-h #`	fastx/y3, tfastx/y3 only - penalty for a frameshift between codons
`-j #`	`fasty3`, `tfasty3` only - penalty for a frameshift against a codon
`-t #`	translation table - `fastx/y3`, `tfastx/y3`, and `tfasta3` now support the BLAST translation tables. See `http://www.ncbi.nlm.nih.gov/htbin-post/Taxonomy/wprintgc?mode=c/`
`-y #`	Width for band optimization; by default 16 for DNA and protein *ktup= 2*; 32 for protein *ktup= 1*
`ssearch3` command line options	
`-f #`	penalty for first residue in a gap
`-g #`	penalty for additional residues in a gap

Fig. 4 goes near here.

Command line options can be divided into five general categories: (1) scoring parameter options, (2) statistics options, (3) algorithm-specific options, (4) file specification options, and (5) output options.

4.1 Changing the scoring parameters

All the programs in the FASTA3 package calculate sequence alignments using two types of scoring parameters: a substitution matrix and gap penalties. The default scoring matrix, gap penalties, E() value cutoff, and comparison algorithm are shown in Table 12. The `fasta3`, `ssearch3`, `fastx/y3` and `tfastx/y3` programs use the BLOSUM50 scoring matrix (16) for protein sequence (and translated protein sequence) comparisons. Alternate protein scoring matrices can be specified with the -s option. Available protein matrices include BLOSUM62 (`-s BL62`) and BLOSUM80 (`-s BL80`), PAM250 (`-s P250`) and PAM120 (`-s P120`) (17, 18), and low evolutionary distance matrices MDM10 (`-s M10`) and MDM20 (`-s M20`) (19). In addition, any scoring matrix can be used by providing a file name for the file containing the substitution values (`-s matrix.file`). Version 3 of the FASTA programs uses the same substitution matrix format as the `blastp` programs, and the `pam` program distributed with the BLAST package can be used to generate appropriately formatted matrices.

Table 12: FASTA Program Defaults

program	query	library	scoring(-s)	gap (-f, -g)	frameshift	-E()	alignment

			matrix	penalties	(-h,-j)	cutoff	
`fasta3`	protein	protein	BLOSUM50	-12/-2		10.0	Smith-Waterman
	DNA (1 strand)	DNA	+5/-4	-16/-4		2.0	band Smith-Waterman[a]
`ssearch3`	protein	protein	BLOSUM50	-12/-2		10.0	Smith-Waterman
	DNA (1 strand)	DNA	+5/-4	-16/-4		2.0	Smith-Waterman
`fastx3`	DNA (1 strand)	protein	BLOSUM50	-15/-2	-20	5.0	Smith-Waterman[b]
`fasty3`	DNA (1 strand)	protein	BLOSUM50	-15/-2	-20/-20	5.0	Smith-Waterman[b]
`tfastx3`	protein	DNA	BLOSUM50	-15/-2	-20	5.0	Smith-Waterman[b]
`tfasty3`	protein	DNA	BLOSUM50	-15/-2	-20/-20	5.0	Smith-Waterman[b]
`fastf3`	mixed peptides	protein	MDM20			5.0	
`tfastf3`	mixed peptides	DNA	MDM10			5.0	

[a] ref. 28; [b] ref. 15

For DNA sequence comparisons, the substitution matrix scores +5 for a match and -4 for a mismatch (+2 for match to an ambiguous nucleotide, -1 for a mismatch to an ambiguous residue). Alternate DNA substitution matrices can be specified using the -s dna-matrix.file option.

The BLOSUM50 matrix works well for recognizing very distant relationships (and works well for long, closely related sequences as well). Searches with short sequences (18) or for closely related sequences (e.g. mouse proteins against mouse ESTs) will be more effective with "shallower" scoring matrices—matrices like MDM10 and MDM20 that are optimum for small amounts of change in very short sequences.

Gap penalties in the FASTA programs can be changed with the -f and -g options; -f specifies the cost of the first residue in a gap and -g specifies the cost of each additional residue. An alternate representation of gap penalties takes the form: $q + rk$, where q is the penalty for opening a gap and r is the penalty for each residue in the gap (k is the length of the gap). Thus, `-f -12`, `-g -2` (the default for protein searches) is equivalent to: $q = 10$, $r = 2$. Protein substitution matrices like BLOSUM50 and PAM250, which are scaled in 1/3-bit units (18), work well with gap penalties of `-12/-2` or `-14/-2` (20), while scoring matrices like BLOSUM62 and PAM120, which are scaled in 1/2-bit units, work well with a lower initial residue penalty, (`-f -8`).

Just as "shallower" substitution matrices may be appropriate for comparisons between closely related sequences (e.g. mammals), higher gap penalties may be appropriate as well.

Using a MDM20 scoring matrix with gap penalties of -20/-4 will cause the program to recognize, with very high expectation values, sequences that have diverged by about 20-40%, but the program will probably miss clear homologs that share less than 30% protein sequence identity.

The `fastx3`/`tfastx3` and `fasty3`/`tfasty3` programs provide additional gap parameters. `fastx3`/`tfastx3` uses `-h` to specify the cost of a frameshift (which must, because of the nature of the `fastx3` algorithm, fall between two codons). `fasty3`/`tfasty3` uses `-h` to set the cost of a between-codon frameshift and `-j` to specify the cost of a frameshift that within a codon. When searching with EST sequences that contain approximately 5% errors, the default values `-h -20` and `-j -20` work well (15). However, if the DNA sequences are known to be relatively error free, searches with higher frameshift penalties are appropriate, as they will reduce the noise from out-of-frame alignments.

In general, the default gap parameters provided by the FASTA programs are at the lower end of the useful range. Reducing the gap penalties more will often cause alignments to shift from local to global, and thus violate the assumptions underlying the statistical estimates. Small increases in the initial residue (`-f`) penalty will sometimes slightly improve the expectation value of an alignment, but researchers should be suspicious of borderline scores that change dramatically with different gap penalties. Changes in substitution matrices usually have a greater effect than small changes in gap penalties; the expectation values from searches with the PAM250 matrix are often 10^{-3}–10^{-10} lower than when BLOSUM50 is used. For example, for the scores shown in Table 7, the E()-values for the alignments of `gtt1_drome` and `xuzm32`, `xuzm31`, and `xuzm1` drop from 8.5×10^{-8}, 2.5×10^{-6}, and 8.8×10^{-5} to 7.1×10^{-5}, 0.001, and 0.15 when the PAM250 matrix is used. When evaluating the significance of an alignment using the Monte-Carlo `prss3` program, one should be certain to use the same substitution matrix and gap penalties.

4.2 Alternate statistical estimates

One of the strengths of the FASTA3 package is its ability to estimate accurately the statistical significance of a local similarity score, regardless of whether it was calculated from a protein:protein, DNA:DNA, or protein:translated-DNA alignment. The programs in the FASTA3 package calculate expectation values based on parameters estimated from the distribution of scores from "unrelated" sequences. Thus, the statistical estimates are accurate for the typical case of a search against a database containing tens of thousands of unrelated sequences, but they will not be accurate if the database does not contain unrelated sequences. The FASTA3 programs provide six statistical estimation options (Table 13; ref. 6). The `-z 3` option is of particular interest, as it can be used when searching databases that do not contain unrelated sequences, or even when comparing a pair of sequences.

Table 13: Statistics options

`-z –1`	No statistical estimates. Sometimes necessary when there are no unrelated sequences in the database.
`-z 0`	Unscaled statistical estimates. Estimates are calculated from the mean and and variance of the sequence similarity scores. Typically used when all of the library sequences have about the same length.

`-z 1`	Regression-scaled estimates. Mean and variance of the similarity scores are calculated after correcting the scores for a log(n) effect.
`-z 2`	Log-corrected estimates. Provided for historical purposes only; this method is out of date and should not be used.
`-z 3`	Altschul-Gish estimates (protein only). Instead of estimating the parameters from the data, pre-calculated parameters published by Altschul and Gish (29) are used. `-z 3` is the only option for estimating the significance of an alignment when unrelated sequences are not the majority of the searched library.
`-z 4`	An alternative to `-z 1` that uses a different method for removing high scoring, potentially related sequences during the parameter estimating process.
`-z 5`	An alternative the `-z 1` that also uses regression of the score variance with log(n) (library sequence length). While `-z 5` is likely to provide somewhat more accurate estimates than `-z 1`, it is also more sensitive to problems with the data, particularly when relatively small libraries (< 500 entries) are searched.

The dependence of statistical significance on database size can complicate comparisons of searches on different databases. The "`-Z number`" option can be used to force the program to pretend that a database of size "number" was searched, e.g. "`-Z 100000`" might be used to reflect the consensus that there are ~100,000 mammalian genes. ("number" should never be smaller than the actual size of the database searched.) This option is particularly important in combination with –z 3 when searching a small set of pre-selected sequences.

4.3 Input options

The FASTA programs provide a number of options that change how the query sequence is used and how the database is selected (Table 14). The most commonly used input option is `-i`, which causes a DNA search to use the reverse complement of the query sequence. (Unlike BLASTN and the GCG version of FASTA, the U. of Virginia FASTA programs do not automatically search with both the forward and reverse DNA strands when a DNA query is used.)

Table 14: Input options

`@`	In addition to using file names, the FASTA3 programs can accept query sequences from the stdin file stream on Unix and Windows computers. In this case, all information must be given on the command line, e.g.: `fasta3 -q @ /slib/swiss.seq 1 < query.aa` indicates that the input will come from `stdin (< query.aa)` and that the `swiss.seq` library will be searched with *ktup=1*. The @ option is most commonly used with `perl` scripts on WWW servers.
`:#-#`	Specify a sub-sequence. Query sequence file names can be followed by a ":" and a range of numbers to specify a portion of a sequence. If the first number is not given, 1 is assumed. If the last number is not given, the subsequence extends to the end of the sequence. Thus, `gtt1_drome.aa:51-150` specifies the 100 residues beginning at residue 51. Subsequence ranges can be given when the query sequence is entered on the command line or when prompted by the program. They can also be entered

	after an "@" (stdin) symbol. Subsequence ranges can only be used for the first (query sequence).
`-i`	(DNA queries only) Search with the reverse complement of the query sequence.
`-l file`	Identify the FASTLIBS file used to locate sequence databases.
`-n`	Force the input (query) sequence to be read as DNA (`fasta3` and `ssearch3` only).
`-N #`	Read long library sequences (such as bacterial genomes) in chunks of "#" residues; e.g. -N 5000 would read long sequences in 5000 residue portions.
`-q/Q`	Quiet. Do not prompt for input.

The FASTA programs make it easy to specify a search with only part of the query sequence with the ":" modifier to the query sequence file name. The command:

```
fasta3 gtt1_drome.aa:1-100 s
```

searches the database specified by the "`s`" abbreviation with the first 100 residues of the query sequence `gtt1_drome`.

`fasta3` and `ssearch3` use a simple algorithm to decide if a query sequence is likely to be protein or DNA. If the sequence is more than 85% A+C+G+T, it is assumed to be DNA; otherwise it is treated as a protein sequence. The `-n` option forces a query sequence to be treated as DNA; the `-n` option is required for DNA sequences provided through the `stdin` (`@`) option (Table 14). Unlike the BLAST programs, the FASTA programs currently report only the best alignment between the query sequence and the library sequence, even when the library sequence is very long and may contain hundreds of genes. By default, FASTA breaks up long DNA sequences into ~80,000 nucleotide pieces, but this size is too large for gene dense bacterial, yeast, and C. elegans genomes. The `-N 5000` option tells `fasta3` and `tfastx/y3` to read long DNA sequences in chunks of 5000 nucleotides. This is essential when scanning large, gene dense DNA sequences.

Table 15: Output options

`-a`	(`fasta3` and `ssearch3` only) show the query and library sequences in their entirety, not just the portion that aligns.
`-A`	(`fasta3` DNA only) `fasta3` does a full Smith-Waterman (22) alignment for protein sequences (and translated `fastx/y3` and `tfastx/y3` alignments) but only a band-limited alignment for DNA:DNA alignments. The -A option forces `fasta3` to do a full Smith- Waterman alignment for DNA sequences. This can slow the program down substantially if one of the sequences is quite long.
`-b #`	The number of high-scoring library sequences scores to be shown.
`-d #`	The number of high-scoring alignments to be shown.
`-E #`	The expectation (E()) value cutoff for showing scores and alignments. By default, `-E 10` for protein:protein comparisons, `-E 5` for translated DNA:protein comparisons, and `-E 2` for DNA:DNA comparisons. The -E cutoff overrides the `-b` and `-d` options; to ensure that at least 20 scores and 5 alignments are shown, the options: `-E 1000.0 -b 20 -d 10` would be used.
`-F #`	A lower-bound expectation value cutoff that prevents very closely related sequences from being shown. `-F 1e-4` will prevent the programs from

	showing library sequences with E() <10^{-4} . This option is useful for focussing on distant homologues in large protein families with many close homologues.
`-H`	Do not show the histogram.
`-L`	Provide long sequence descriptions with the alignment. Some sequence library formats (particularly reformatted GCG libraries) include a lot of uninformative text before the actual sequence description. With the `-L` option, all the sequence description available is displayed with the alignment.
`-m #`	See Table 16.
`-O file`	Send results to "file". Unix and Windows users should use the "> file" method for output redirection.
`-r file`	Send intermediate results for all sequences to "file".
`-w #`	Width of alignment output. The FASTA programs display alignments with 60 residues per line by default; this width can be increased to 200 residues with the -w option.
`-W #`	Amount of sequence context. `fasta3` and `ssearch3` provide neighboring sequence context in the alignment (translated fastx/y3 and tfastx/y3 do not). The amount of context is typically one half of an output line, but this amount can be increased or reduced with the `-W` option.
`-x "# #"`	Sequence coordinates. Normally, the FASTA programs assume that each sequence begins at residue 1. On occasion, it is useful to use a different initial coordinate, such as when comparing a cDNA to the encoding gene or when working with only a portion of a sequence. -x “1 -751” would tell `fasta3` to begin the numbering of the library sequence at “-751” rather than “1”. On Unix, DOS, and Macintosh systems, the two numbers must be surrounded by double quotation ("...") marks.

4.4 Changing the output appearance

Many of the FASTA command line options change the appearance of the alignment output (Table 15). Options are available to change the number of residues displayed on an alignment line, to change the numbering of the residues, and to change the format of the alignment. Two options are of particular interest: `-m 5` provides both the sequence alignment and a crude graphical mapping of the aligned region against the query sequence. This graph makes it much easier to see quickly the parts of the query that align with the different library sequences, and thus can highlight query sequences with separable domains. The `–m 6` option is identical to `-m 5`, but provides `html` mark up commands and links to Entrez and other sites for re-searching to confirm relationships with the library sequence.

Figure 4: Alternative output formats

A.

```
>>GTT1_MUSDO GLUTATHIONE S-TRANSFERASE 1 (EC 2.5.1.18) (C (208 aa)
 initn: 1229 init1: 1229 opt: 1230 Z-score: 1472.4 expect() 2.3e-75
Smith-Waterman score: 1230;  85.024% identity in 207 aa overlap

               10        20        30        40        50        60
gi|121 MVDFYYLPGSSPCRSVIMTAKAVGVELNKKLLNLQAGEHLKPEFLKINPQHTIPTLVDNG
       .:::::::.:::::.:::::.::::::::::::::::::::::::::::::::::::.
GTT1_M  MDFYYLPGSAPCRSVLMTAKALGIELNKKLLNLQAGEHLKPEFLKINPQHTIPTLVDGD
                10        20        30        40        50
```

B.

```
>>GTT1_MUSDO GLUTATHIONE S-TRANSFERASE 1 (EC 2.5.1.18) (CLASS-THETA).    (208 aa)
 initn: 1229 init1: 1229 opt: 1230 Z-score: 1615.1 expect() 2.6e-83
Smith-Waterman score: 1230;  85.024% identity in 207 aa overlap

               10        20        30        40        50        60        70
gi|121 MVDFYYLPGSSPCRSVIMTAKAVGVELNKKLLNLQAGEHLKPEFLKINPQHTIPTLVDNGFALWESRAIQVYLVE
       x        x     x     x x                                   xX        x
GTT1_M  MDFYYLPGSAPCRSVLMTAKALGIELNKKLLNLQAGEHLKPEFLKINPQHTIPTLVDGDFALWESRAIMVYLVE
                10        20        30        40        50        60        70
```

Fig. 4: Alternative output formats—Alignments of `gtt1_drome` with `gtt1_musdo` are shown using the default (A) program parameters and (B) the command line options:

```
-f –14 -w 75 –L –m 1
```

(see text for details).

Table 16: Alignment options

`-m 0`	Highlight identical aligned residues with ":", conservative replacements with "."
`-m 1`	Identities are not highlighted. Highlight conservative replacements with "x", non-conservative replacements with "X".
`-m 2`	Highlight identities with ".", non-identical residues with the residue.
`-m 3`	The alignments are printed as two fasta format sequence entries with "-" indicating gaps. These files are sometimes useful as input to other programs.
`-m 4`	Do not show an alignment; show a graph (------) of where the aligned region maps onto the query sequence. Useful for highlighting different domains in proteins.
`-m 5`	A combination of -m 0 and -m 4 that shows both the mapping and the alignment.
`-m 6`	Similar to -m 5, but includes html commands for a WWW browser like Netscape or Internet Explorer and links to simplify looking up the library sequence and re-searching the database.
`-m 10`	Parseable output designed to be read by other computer programs. Each alignment is a series of labeled tags that specify the beginning, end, score, search parameters, and other information.

5. BEYOND SEQUENCE HOMLOGY—IDENTIFYING NEW PARALOGS

The use of the FASTA and BLAST programs for identifying distantly related sequences has been extensively reviewed (3-5), so in this last section we will consider a slightly different problem that exploits the flexibility of the FASTA programs and the high quality of their alignments.

Here, we seek to identify new paralogs of known human or mouse families from EST databases. For example, two human prostaglandin synthase enzymes are known, COX1 (`pgh1_human`) and COX2 (`pgh2_human`), in humans, mice, rats, and other mammals. Prostaglandin synthases are targets of non-steroidal anti-inflamatory drugs, including aspirin and ibuprofen. Thus, there is great interest in finding additional members of this family and it is certainly possible that additional prostaglandin synthases have been sequenced, either by large scale EST sequencing or by genomic sequencing.

5.1 Overall strategy

Paralogs are members of a gene family (and are thus related or homologous) that differ from other sequences in the family because of gene duplication events. (Orthologous genes differ because they are found in different species.) A search of the SwissProt database (Table 17) shows the two prostaglandin synthase (PGH) subfamilies, but also shows distantly related peroxidases. The human PGH1 and PGH2 isoenzymes share about 65% sequence identity (E() $< 10^{-165}$). (In contrast, orthologous human and mouse PGH1 sequences share 89.3% identity.) We expect a new human PGH synthase to share very strong similarity to PGH1 and PGH2 (E() $<10^{-20}$) but to share less than 80% identity to either PGH1 or PGH2. Since we will be scanning EST databases to find the new paralogs, we expect that sequences with > 90-95% identity are probably from mRNAs for known proteins that have sequencing errors, but that sequences that are 50-90% identical are candidate paralogs.

Table 17: Prostaglandin synthase search results

The best scores are:		len	E(74357)
PGH1_HUMAN	prostaglandin G/H synthase 1	599	3.9e-264
PGH1_SHEEP	prostaglandin G/H synthase 1	600	2.3e-244
PGH1_MOUSE	prostaglandin G/H synthase 1	602	9.5e-237
PGH2_CHICK	prostaglandin G/H synthase 2	603	1.2e-168
PGH2_HUMAN	prostaglandin G/H synthase 2	604	1.9e-165
PGH2_MOUSE	prostaglandin G/H synthase 2	604	2.4e-164
PGH2_CAVPO	prostaglandin G/H synthase 2	604	1.7e-163
PGH2_RAT	prostaglandin G/H synthase 2	604	1.4e-162
PERM_MOUSE	myeloperoxidase prec.	718	0.0001
PERO_DROME	peroxidase prec.	690	0.00024
PERT_HUMAN	thyroid peroxidase prec.	933	0.0003
PERM_HUMAN	myeloperoxidase prec.	745	0.00034
PERT_PIG	thyroid peroxidase prec.	926	0.0029
PERL_BOVIN	lactoperoxidase prec.	712	0.016
PERT_MOUSE	thyroid peroxidase prec.	914	0.02
PERL_HUMAN	lactoperoxidase LPO	324	0.027
PERT_RAT	thyroid peroxidase prec.	914	0.089
FBP1_STRPU	fibropellin I prec.	1064	0.16

PGCN_RAT	neurocan core prot. prec.	1257	0.21
FBP3_STRPU	fibropellin C prec.	570	0.31
PGCN_MOUSE	neurocan core prot. prec.	1268	0.33
PERE_MOUSE	eosinophil peroxidase prec.	716	0.51
NOTC_DROME	neurogenic locus notch prot.	2703	0.74
DLK_MOUSE	delta-like prot. prec.	385	0.86
PERE_HUMAN	eosinophil peroxidase prec.	715	0.92
NTC1_MOUSE	neurogenic locus notch homolog	2531	0.94

Results of a `fasta3` (*ktup=2*) search with `pgh1_human` against the Swissprot protein sequence database.

To identify new `pgh1_human` paralogs, we will search the human EST database (obtained from ftp://ncbi.nlm.nih.gov/blast/db/) with the `pgh1_human` and `pgh2_human` protein sequences using the `tfasty3` program. `tfasty3` is used because: (1) we wish to compare a protein query to a DNA (EST) database; and (2) we will use both the expectation value E() and the percent identity to characterize matches, so a high-quality protein:DNA alignment is required (`tfastx3` is faster but produces a lower quality alignment, ref. 15). We will then examine the EST sequences that share significant similarity and categorize them as orthologous to `pgh1_human`, `pgh2_human`, or a new paralog.

5.2 Statistical significance and percent identity

While our goal is to identify sequences that are similar to, but not identical with, known prostaglandin synthases, conventional similarity criteria (E()-value and percent identity) do not fully capture the information we seek. As the results of the `pgh1_human` and `pgh2_human` `tfasty3` searches demonstrate (Table 18), EST sequences that share higher sequence identity do not necessarily have better E()-values.

The discrepancy between E()-value and percent identity reflects the dependence of E()-value on alignment length. EST sequences tend to be partial, so that an orthologous 100% match to the C-terminal 30 amino acids in `gb|N79146` can have a worse expectation value (2.9×10^{-6}) than a 59% identity to a paralogous gene ($E() < 6.7 \times 10^{-19}$). However, percent identity is a poor criterion for similarity, because unrelated sequences (e.g. `gb|AA485017`) can share high identity (66.1% over 62 codons) that does not produce a statistically significant similarity score. Nevertheless, for sequences that share significant similarity, percent identity is a useful measure of sequence difference. Thus, among the statistically significant matches in Table 18, orthologous matches always had percent identities > 90%, with one possible exception (`gb|AA223896`, see below).

Table 18: Prostaglandin synthase ESTs

pgh1_human:		len	[f/r]	opt	E(10^6)	%ident.	I/II
gb|R96180	Pineal_gland_N3HPG	355	[f]	654	3e-38	98.0	I
gb|AA296431	Umbilical vein endothelial	279	[f]	380	6.7e-19	59.1	II
gb|T29235	Human Bone	257	[f]	358	2.2e-17	63.3	II
gb|AA037294	Senescent_fibroblasts_NbHSF	471	[f]	304	3.1e-13	98.0	I
gb|AI022012	Senescent_fibroblasts_NbHSF	537	[r]	248	3.5e-09	64.5	II
gb|N79146	Multiple_sclerosis_2NbHMSP	544	[f]	207	2.9e-06	100.0	I
gb|AA223896	NT2 neuronal precursor	97	[f]	185	1.3e-05	80.0	??

gb|AA485017	NCI_CGAP_GCB1	208	[f]	124	0.72	66.1	
pgh2_human:		len	[f/r]	opt	E(10^6)	%ident.	I/II
gb|AA296431	Umbilical vein endothelial	279	[f]	574	1.4e-35	96.8	II
gb|T29235	Human Bone	257	[f]	536	1e-32	92.9	II
gb|AI022012	Senescent_fibroblasts_NbHSF	537	[r]	541	1.1e-32	95.8	II
gb|R96180	Pineal_gland_N3HPG	355	[f]	410	6.3e-23	65.8	I
gb|AA223896	NT2 neuronal precursor	97	[f]	136	0.01	50.0	??
gb|AA885610	NCI_CGAP_Lu5	320	[f]	141	0.018	46.3	
gb|AA911293	NCI_CGAP_Lu5	172	[f]	131	0.049	43.6	

Results from searches with pgh1_human and pgh2_human against the BLAST est_human database using tfasty3 and with the default BLOSUM50 scoring matrix. pgh1 (COXI) or pgh2 (COXII) orthologs are labeled in the right column.

5.3 Shifting evolutionary horizons with scoring matrices

Examination of the high scoring ESTs found with pgh1_human and pgh2_human in Table 18 suggests that all but one of the ESTs share > 90% identity with either pgh1_human or pgh2_human. The exception, gb|AA223896, shares only 80% identity with pgh1_human and 50% identity with pgh2_human, and thus is a candidate novel paralog prostaglandin synthase.

However, the gb|AA223896 EST sequence is very short (97 nucleotides), and there are only 6 mismatches, half of which are within 20 nucleotides of one end of the sequence. Thus, we must consider whether this is truly a novel paralog, or simply a short, poor-quality sequence of a pgh1_human mRNA that has several errors at one end (as is expected with high-throughput EST sequencing). While the end-sequence error problem could be reduced by ad hoc changes to the alignment code that down-weighted end-mismatches, a simpler approach is to use shallower scoring matrices.

Table 19: Searching with “shallow” scoring matrices

pgh1_human:	len	E(BL50)	%	E(M20)	%	E(M10)	%	I/II
gb|R96180	355	3e-38	98.0	2.3e-72	99.0	6.5e-75	100.0	I
gb|AA296431	279	6.7e-19	59.1	6.8e-25	61.3	1.3e-22	62.4	II
gb|T29235	257	2.2e-17	63.3	5.3e-22	64.8	2.6e-18	66.2	II
gb|AA037294	471	3.1e-13	98.0	3e-30	98.0	3.3e-31	97.8	I
gb|AI022012	537	3.5e-09	64.5	1.2e-15	58.8	3.4e-13	60.8	II
gb|N79146	544	2.9e-06	100.0	2.6e-16	100.0	3.0e-17	100.0	I
gb|AA223896	97	1.3e-05	80.0	8.4e-13	87.1	2.8e-12	87.1	??
gb|AA485017	208	0.72	66.1	4.8e-14	84.7	4.1e-14	88.9	??
pgh2_human:								
gb|AA296431	279	1.4e-35	96.8	2.2e-69	96.8	8.0e-72	98.9	II
gb|T29235	257	1e-32	92.9	2.9e-61	94.1	9.1e-63	95.2	II
gb|AI022012	537	1.1e-32	95.8	1.6e-68	96.0	1.1e-70	97.0	II
gb|R96180	355	6.3e-23	65.8	1.0e-30	56.9	9.1e-27	60.3	I
gb|AA485017	208	__a	—	2.4e-05	75.6	3.3e-4	79.1	??
gb|AA223896	97	0.01	50.0	0.01	69.0	0.2	79.2	??
gb|AA885610	320	0.018	46.3	—	—	—	—	
gb|AA911293	172	0.049	43.6	—	—	—	—	

[a]E() values indicated as — were >5.0.

Additional searches with very shallow scoring matrices (MDM20 and MDM10, ref. 19; Table 19) show slightly different, potentially more interesting perspectives. When shallower scoring matrices are used, both orthologous and paralogous alignments become mores statistically significant, and, as expected, the percent identities increase ("shallower" scoring matrices give more positive scores to identities and more negative scores to non-conservative replacements). Of greater interest are two sequences `gb|AA223896` and `gb|AA485017`, which show significant similarity with `pgh1_human` with MDM20 and MDM10. Both sequences are tantalizing candidates for new paralogs (as orthologs consistently have percent identities higher than 90% with MDM20. However, the alignments of both sequences show a large number of frameshifts (which do not affect the percent identity calculation), suggesting that these sequences may have percent identities < 90% because of a poor quality sequence, rather than a novel gene.

The last two entries (`gb|AA885610` and `gb|AA911293`) in the `pgh2_human` search shows that shallow scoring matrices can also be used to quickly rule out high scoring unrelated sequences. The expectation values for those two sequences, which were marginally significant (0.018 and 0.049) scores with BLOSUM50 and were not significantly similar to `pgh1_human`, became very high (E() > 5) when MDM20 and MDM10 were used. Thus, shallower scoring matrices can be used to provide a more stringent test for sequence similarity when near-identity is expected for at least one of the query sequences.[4]

6. SUMMARY

The FASTA3 and FASTA2 packages provide a flexible set of sequence comparison programs that are particularly valuable because of their accurate statistical estimates and high-quality alignments. Traditionally, sequence similarity searches have sought to ask one question: "Is my query sequence homologous to anything in the database?" Both FASTA and BLAST can provide reliable answers to this question with their statistical estimates; if the expectation value E() is < 0.001-0.01 and you aren't doing hundreds of searches a day, the answer is probably yes.

In general, the most effective search strategies follow these rules:

1. Whenever possible, compare at the amino acid level, rather than the nucleotide level. Search first with protein sequences (`blastp`, `fasta3`, and `ssearch3`), then with translated DNA sequences (`fastx`, `blastx`), and only at the DNA level as a last resort (Table 5).
2. Search the smallest database that is likely to contain the sequence of interest (but it must contain many unrelated sequences for accurate statistical estimates).
3. Use sequence statistics, rather than percent identity or percent similarity, as your primary criterion for sequence homology.

---

[4]While MDM20 and MDM10 can serve to provide more stringent alignments, they are not the best matrices, because they were built assuming an evolutionary model. More accurate matrices could be derived from looking at large numbers of EST sequencing errors, and building a matrix that was based on a sequencing error model, rather than evolutionary divergence.

4. Check that the statistics are likely to be accurate by looking for the highest scoring unrelated sequence, using prss3 to confirm the expectation, and searching with shuffled copies of the query sequence (randseq, searches with shuffled sequences should have E() ~1.0).
5. Consider searches with different gap penalties and other scoring matrices. Searches with long query sequences against full-length sequence libraries will not change dramatically when BLOSUM62 is used instead of BLOSUM50 (20), or a gap penalty of -14/-2 is used in place of -12/-2. However, shallower or more stringent scoring matrices are more effective at uncovering relationships in partial sequences.(3, 18), and they can be used to sharpen dramatically the scope of the similarity search.

However, as illustrated in the last section, the E() value is only the first step in characterizing a sequence relationship. Once one has confidence that the sequences are homologous, one should look at the sequence alignments and percent identities, particularly when searching with lower quality sequences. When sequence alignments are very short, the alignment should become more significant when a shallower scoring matrix is used, e.g. BLOSUM62 rather than BLOSUM50 (remember to change the gap penalties).

Homology can be reliably inferred from statistically significant similarity. While homology implies common three-dimensional structure, homology need not imply common function. Orthologous sequences usually have similar functions, but paralogous sequences often acquire very different functional roles. Motif databases, such as PROSITE (21), can provide evidence for the conservation of critical functional residues. However, motif identity in the absence of overall sequence similarity is not a reliable indicator of homology.

Acknowledgements

W.R.P. is supported by a grant from the National Library of Medicine (LM04961).

REFERENCES

1. Wilbur, W. J. and Lipman, D. J. (1983) Rapid similarity searches of nucleic acid and protein data banks, *Proc. Natl. Acad. Sci. USA* **80**, 726-730.
2. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sulton, G. G., Blake J. A., Fitzgerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J.-F., Adams, M. D., Reisch, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S. M., Weidman, J. F., Fuhrmann, J. L., Nguyen, D., Utterback, T. R., Kelley, J. M., Peterson, J. D., Sadow, P. W., Hanna, M. C., Cotton, M. D., Roberts, K. M., Hurst, M. A., Kaine, B. P., Borodovsky, M., Klenk, H.-P., Fraser, C. M., Smith, H. O., Woese, C. R., and Venter, J. C. (1996) Complete genome sequence of the methanogenic archaeon, methanococcus jannaschii *Science* **273**, 1058-1073.
3. Altschul, S. F., Boguski, M. S., Gish, W., and Wootton, J. C. (1994) Issues in searching molecular sequence databases *Nature Genet.* **6**, 119-129.
4. Pearson, W. R. (1996) Effective protein sequence comparison *Methods Enzymol.* **266**, 227-258.
5. Pearson, W. R. (1997) Identifying distantly related protein sequences *Comput. Appl. Biosci.* **13**, 325-332.

6. Pearson, W. R. (1998) Empirical statistical estimates for sequence similarity searches *J. Mol. Biol.* **276**, 71-84.
7. Pearson, W. R. and Lipman, D. J. (1988) Improved tools for biological sequence comparison *Proc. Natl. Acad. Sci. USA* **85**, 2444-2448.
8. Lipman, D. J. and Pearson, W. R. (1985) Rapid and sensitive protein similarity searches *Science* **227**, 1435-1441.
9. Bleasby, A. J., Akrigg, D., and Attwood, T. K. (1994) Owl-a non-redundant composite protein sequence database. *Nucleic Acids Res.* **22**, 3574-3577.
10. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) A basic local alignment search tool *J. Mol. Biol.* **215**, 403-410.
11. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped blast and psi-blast: a new generation of protein database search programs *Nucleic Acids Res.* **25**, 3389-3402.
12. Arratia, R., Gordon, L., and Waterman, M. S. (1986) An extreme value theory for sequence matching *Ann. Stat.* **14**, 971-993.
13. Karlin, S. and Altschul, S. F. (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes *Proc. Natl. Acad. Sci. USA* **87**, 2264-2268.
14. Wootton, J. C. and Federhen, S. (1993) Statistics of local complexity in amino acid sequences and sequence databases *Comput. Chem.* **17**, 149-163.
15. Pearson, W. R., Wood, T., Zhang, Z., and Miller, W. (1997) Comparison of DNA sequences with protein sequences *Genomics* **46**, 24-36.
16. Henikoff, S. and Henikoff, J. G. (1992) Amino acid substitutions matrices from protein blocks *Proc. Natl. Acad. Sci. USA* **89**, 10915-10919.
17. Schwartz, R. M. and Dayhoff, M. (1978) Matrices for detecting distant relationships In M. Dayhoff, (ed.), Atlas of Protein Sequence and Structure, volume 5, supplement 3, pp. 353-358 National Biomedical Research Foundation Silver Spring, MD.
18. Altschul, S. F. (1991) Amino acid substitution matrices from an information theoretic perspective *J. Mol. Biol.* **219**, 555-565.
19. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences *Comp. Appl. Biosci.* **8**, 275-282.
20. Pearson, W. R. (1995) Comparison of methods for searching protein sequence databases *Protein Sci.* **4**, 1145-1160.
21. Bairoch, A. (1991) PROSITE: a dictionary of sites and patterns in proteins *Nucleic Acids Res.* **19**, supplement, 2241-2245.
22. Smith, T. F. and Waterman, M. S. (1981) Identification of common molecular subsequences *J. Mol. Biol.* **147**, 195-197.
23. Huang, X. and Miller, W. (1991) A time-efficient, linear-space local similarity algorithm *Adv. Appl. Math.* **12**, 337-357.
24. Waterman, M. S. and Eggert, M. (1987) A new algorithm for best subsequences alignment with application to tRNA-rRNA comparisons *J. Mol. Biol.* **197**, 723-728.
25. Myers, E. W. and Miller, W. (1988) Optimal alignments in linear space *Comp. Appl. Biosci.* **4**, 11-17.
26. Kyte, J. and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein *J. Mol. Biol.* **157**, 105-132.

27. Barker, W. C., Garavelli, J. S., Haft, D. H., Hunt, L. T., Marzec, C. R., Orcutt, B. C., Srinivasarao, G. Y., Yeh, L. S. L., Ledley, R. S., Mewes, H. W., Pfeiffer, F., and Tsugita, A. (1998) The PIR-International protein sequence database *Nucleic Acids Res.* **26**, 27-32.
28. Chao, K.-M., Pearson, W. R., and Miller, W. (1992) Aligning two sequences within a specified diagonal band *Comp. Appl. Biosci.* **8**, 481-487.
29. Altschul, S. F. and Gish, W. (1996) Local alignment statistics *Methods Enzymol.* **266**, 460-480.

Allergy 2001: 56: 478 490
Printed in UK. All rights reserved

ALLERGY
ISSN 0105-4538

Review article

Cross-reactivity of IgE antibodies to allergens

The cross-reactivity of IgE antibodies is of interest for various reasons, three of which are discussed. Firstly, from the clinical view, it is important to know the patterns of cross-reactivity, because they often (but not always) reflect the pattern of clinical sensitivities. We discuss the cross-reactivities associated with sensitization to pollen and vegetable foods: PR-10 (Bet v 1-related), profilin, the cross-reactive carbohydrate determinant (CCD), the recently described isoflavone reductase, and the (still elusive) mugwort allergen that is associated with celery anaphylaxis; cross-reactivities between allergens from invertebrates, particularly tropomyosin, paramyosin, and glutathione S-transferase (GST); and latex-associated cross-reactivities. Clustering cross-reactive allergens may simplify diagnostic procedures and therapeutic regimens. Secondly, IgE cross-reactivity is of interest for its immunologic basis, particularly in relation to the regulation of allergic sensitization: are IgE antibodies to allergens more often cross-reactive than IgG antibodies to "normal" antigens? If so, why? For this discussion, it is relevant to compare not only the structural relation between the two allergens in question, but also the relatedness to the human equivalent (if any) and how the latter influences the immune repertoire. Thirdly, prediction of IgE cross-reactivity is of interest in relation to allergic reactivity to novel foods. Cross-reactivity is a property defined by individual antibodies to individual allergens. Quantitative information (including relative affinity) is required on cross-reactivity in the allergic population and with specific allergens (rather than with whole extracts). Such information is still scarce, but with the increasing availability of purified (usually recombinant) allergens, such quantitative information will soon start to accumulate. It is expected that similarity in short stretches of the linear amino-acid sequence is unlikely to result in relevant cross-reactivity between two proteins unless there is similarity in the protein fold.

R. C. Aalberse, J. H. Akkerdaas, R. van Ree

CLB and Laboratory for Experimental and Clinical Immunology, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

Key words: allergen structure; allergenicity; cross reactive carbohydrate determinant (CCD); cross reactivity.

Prof. R. C. Aalberse
CLB, Department of Immunopathology
Plesmanlaan 125
1066 CX Amsterdam
The Netherlands

Accepted for publication 5 February 2001

A short history of allergen cross-reactivity

Allergen cross-reactivity works usually, but not always, "according to the book": in general, cross-reactivity reflects the phylogenetic relations between organisms. A phylogenetic relation results in a high degree of homology in the primary structure of the proteins (the amino-acid sequence; see Table 1 for the database accession codes of the proteins that will be discussed). High homology in the primary sequence results in homologous 3-D structures and thus, potentially, in cross-reactivity. Serum albumins from vertebrates are often cross-reactive (1). Homologous allergens from phylogenetically related grasses tend to be cross-reactive (2), and IgE antibodies to allergens from peanuts are often cross-reactive with homologous proteins in soybean and other legumes.

However, cross-reactivities between much more distantly related organisms have been known for some time. Some of the highlights in the history of unexpected cross-reactions in allergy are ragweed/banana, which was recognized as early as 1970 (3), birch/apple (4), cross-reacting carbohydrate determinant (5), profilin (6, 7), mugwort/celergy (8), latex/banana/avocado (9), and mite/snail/shrimp (10–14).

The birch-apple story started in 1977 when Hannuksela & Lahti (4) described the association on clinical grounds. For two independent reasons, cross-reactivity was not easily demonstrable at that time. Firstly, the apple extracts that were then available were deficient in the relevant proteins. Only when Björkstén et al. adapted an extraction procedure (15) could suitable extracts be prepared. Secondly, the apple extracts contained cross-reacting structures (CCDs [5]), but these did not show well in the technique that was used at that time: crossed radioimmunoelectrophoresis (16, 17). This changed dramatically in the early 1980s when immunoblotting became the preferred procedure. This resulted in the recognition of two types of cross-reactivity, one due to proteins and the other due to sugars (glycans on glycoproteins); the latter type was

called cross-reacting carbohydrate determinant (CCD) (5). Soon it became clear that, in addition to CCD-type cross-reactivity, there had to be at least two proteins involved in the birch-apple cross-reactivity, one being the major birch-pollen allergen and the other an elusive, slightly smaller protein that was even more broadly cross-reactive, as with grass pollen (18). This protein was identified by molecular cloning and proved to be the proline-binding protein, profilin (6). Thus, the molecular-biology approach not only led to the identification of this most important cross-reacting allergen, but it also provided an efficient purification procedure: affinity chromatography with polyproline.

Two more allergens have recently been added to this list of pollen-associated food allergens: one is the homolog of the sweet-tasting 31-kDa protein thaumatin (PR-5), which has been described in apple (19), cherry (20), and strawberry among other sources. The other is isoflavone reductase (21), a 33–35-kDa protein.

The mugwort-celery connection (8, 22) has been investigated extensively by Wüthrich and coworkers (23–28) and Vallier et al. (29, 30). The molecular basis for this association between mugwort pollen sensitivity and anaphylactic reactivity to a heat-stable allergen in celery (24, 28, 29, 31–33) is still unresolved. Part of the serologic cross-reactivity is due to profilin (30) and CCD, but this does not explain the striking clinical phenotype. Mugwort pollen does not contain the PR-10 protein (a Bet v 1 homolog), and so this is not a candidate either (34). One potential candidate is a lipid transfer protein (LTP) (see below) (35). Another candidate is the major allergen Art v 1 (36), which is heavily glycosylated. This may hamper the expression of an immunologically fully reactive (eagerly awaited) recombinant protein that might be used for analyzing this intriguing and clinically highly relevant cross-reactive system.

LTPs, which are also classified as PR proteins (PR-14) (37), are small, very stable, proteins. They have been identified as the major cause of anaphylactic sensitivity to fruits (38–40). Cross-reactivity has been found among LTPs from Rosaceae species (apple/almond/peach). Based on the amino-acid identity data (less than 35% identity between LTP from mugwort and LTP from *Parietaria*), cross-reactivity among LTPs is likely to be lower than among, for example, profilins. Our own preliminary data suggest significant cross-reactivity between LTPs from celery and apple.

The relations within the latex/fruit cluster (9, 37, 41, 42) are even more complex than in the previous clusters. Exposure to rubber latex occurs by at least three routes: via the airways (as dust mostly from gloves in medical settings), via the skin, and parenterally during operations and other procedures. In each of these three situations, the allergens involved may differ. Reactivity to foods associated with latex sensitivity, such as avocado and banana, occurs with very different prevalences. Last but not least, from the point of view of a protein chemist, rubber latex is a more challenging biologic fluid than a pollen extract.

Chitinase-related proteins (the basic 25–35-kDa variants, PR-3 and PR-4 [37]) are a major source of cross-reactivity between rubber latex (Heb v 6 [43]) and fruits such as banana, avocado (Pers a 1 [44]), and chestnut. Hevein, which is the smaller chitin-binding N-terminal domain in the precursor, is easily severed. It appears to be the major cross-reactive part (45). The remaining part of the precursor protein is related to the wound-induced proteins (Win-proteins). These are also well conserved, but do not appear to contribute to cross-reactivity. Because these enzymes often cleave not only chitin (poly-*N*-acetylglucosamine) but also a structurally related bacterial cell-wall polymer, which consists of *N*-acetylglucosamine and *N*-acetylmuramic acid, they are also referred to as lysozymes. Another chitinase-related protein in rubber latex, hevamin, does not appear to be important in cross-reactivity.

In addition to the chitinases, rubber latex contains a number of other proteins that may to some degree contribute to cross-reactivity. There is a discrepancy regarding the cross-reactivity of Hev b 7, a 46-kDa protein (46) related to the potato allergen patatin (47, 48). Seppälä et al. (49) reported cross-reactivity in 10/35 adult patients (but not in children). However, in a similar patient population, no cross-reactivity was found by Sowka et al. (50).

Invertebrate allergens

In addition to the above-mentioned examples of cross-reactivity between plant-derived proteins, there are some examples of cross-reactivity between allergens from invertebrates, particularly between mite and snail (11, 12); mite, shrimp, and cockroach (13); mite and caddis fly (51); and mites and schistosomes. Three of the proteins that are involved in these examples are tropomyosin, paramyosin (52), and glutathione S-transferase (GST).

The CCD structure (or, rather, structures, because this is really a family rather than one specific structure) has been largely determined, both chemically (53–70) and in 3-D (71, 72) (Fig. 1). This structural information refers to the N-linked glycans. One important unresolved question is whether and how O-linked sugars fit into this picture. Another, even more important, question is the relation (or, rather, lack of relation) between IgE binding and biologic symptoms (73). In most cases, the biologic activity is low compared to the immunochemical activity, but in several cases biologic activity was not low at all (74). In this context, the data on the sea-squirt glycan are also of interest (75–78). One possible factor might be the source of the glycan used for testing. If the glycan is derived from a source material that caused symptoms in the patients, the

Table 1. Entry codes for proteins referred to in figures. Sequences are from Swiss Prot database (sp) or from trEMBL database (tr). These can be accessed at http://www.expasy.ch

Amb e 1	sp	P27759
Avocado chitinase	tr	P93680
Banana glucanase	tr	O22317
Banana pectate lyase	tr	Q9SDW4
Barley group 5	tr	O04828
Bet v 1	sp	P15494
Bet v 4	tr	O82040
Birch chitinase	tr	Q9M3T4
Birch isoflavone reductase	tr	O65002
Birch profilin	sp	P25816
Bovine albumin	sp	P02769
Bovine lactoglobulin	sp	P02754
Brazil nut 2S	sp	P04403
Cauliflower LTP	sp	Q42641
Cedar Cry j 1	sp	P18632
Chitinase avocado	tr	P93680
Chitinase potato	sp	P52403
Cladosporium enolase	sp	P42040
Cod parvalbumin	sp	P02622
Cucumber patatin	tr	O23784
Cyn d 1	tr	O04701
Cyn d profilin	sp	O04725
Dau c 1	sp	O04298
Der f 1	sp	P16311
Der f 2	sp	Q00855
Der p 1	sp	P08176
Der p 2	sp	P49278
Dog serum albumin	tr	Q9TSZ4
Equ c 1	sp	Q95182
Extensin	sp	Q03211
Fruit fly paramyosin	tr	O18392
Grapefruit isoflavone reductase	tr	O49820
Hev b 7	tr	O65811
Hevea chitinase	sp	P23472
Hevea enolase	tr	Q9LEJ0
Hevea glucanase (Hev b 2)	sp	P52407
Hevein avocado	tr	P93680
Hevein potato	sp	P52403
Horse Equ c 1	sp	Q95182
Human A1micro	sp	P02760
Human albumin	sp	P02768
Human calmodulin	sp	P27482
Human cathepsin K	sp	P43235
Human enolase	sp	P06733
Human GST	sp	P46439
Human myosin	tr	Q9Y622
Human parvalbumin	sp	P20472
Human profilin	sp	P35080
Human serpin SCCI	sp	P29508
Human tropomyosin	sp	P06753
Jun a 3	sp	P81295
Jun o 2	tr	O64943
Jun o 2	tr	O64943
Lol p 1	sp	P14946
Lol p 5	sp	Q40240
Maize vicilin	tr	Q03865
Mal d 1	tr	Q43549
Mite GST (Der p 8)	sp	P46419
Mite tropomyosin	tr	O16188
MnSOD Asp	tr	Q9SM64
MnSOD Human	sp	P04179
MnSOD Peach	tr	Q9P945
Mustard 2S	sp	P15322
Ole e 1	sp	P19963
Orange glucanase	tr	O23783
Ovalbumin	sp	P01012
Par j 1	sp	P43217
Peach LTP	sp	P81402
Peanut vicilin (Ara h 1)	sp	P43237
Pear isoflavone reductase	tr	O81355
Phl p 1	sp	P43213
Phl p 5	tr	O81341
Phl p 7	tr	O04131
Pig serpin	sp	P80229
Potato patatin	sp	P15478
Rat n 1	sp	P02761
Salmon parvalbumin	sp	Q91483
Schisto GST	sp	P35661
Schisto paramyosin	sp	P06198
Sesame 2S	tr	Q9XHP1
Shrimp tropomyosin	sp	Q25456
Soy vicilin	tr	O22121
Strawberry thaumatin	tr	Q9SBT2
Tomato calmodulin	sp	P27161
Tyr p 2	sp	O02380

chance of a biologic effect seems to be higher. This suggests that there may be some subtle differences in structure that are not obvious with the current (immuno-) chemical analyses. It is likely that other families of glycans exist with similarly broad cross-reactivity patterns – for example, mold/yeast glycoconjugates.

This historical survey beautifully illustrates the need for close collaboration between clinicians and the laboratory. The observation of an association between sensitivities to allergens is an important starting point. The distinction between cosensitization and cross-reactivity requires *in vitro* experiments. Characterization of the cross-reacting component by immunoblotting, protein fractionation, and recombinant DNA technology requires feedback from the clinician, because some allergens that seem to be important in the laboratory prove to be (almost) without clinical relevance. If a patient is clinically reactive to two allergen source materials and the patient's serum contains a strongly cross-reactive antibody, it is tempting to assume that these two observations are causally related. This assumption has quite often proved to be wrong (if so, the confusion is usually due to the coincidental presence of IgE to CCD, which shows strongly on immunoblots, but has little *in vivo* activity).

Diagnostic and therapeutic issues

If two allergens are very similar, it does not increase the diagnostic accuracy to include both in the diagnostic test panel (79, 80). Similarly, successful treatment with one allergen is likely to relieve symptoms the other allergen as well (81). Particularly in the therapeutic situation, it is likely to be relevant to identify the primary sensitizing allergen, as this will cover the widest spectrum of specificities. Such identification can be made by *in vitro* reciprocal inhibition tests. Cross-

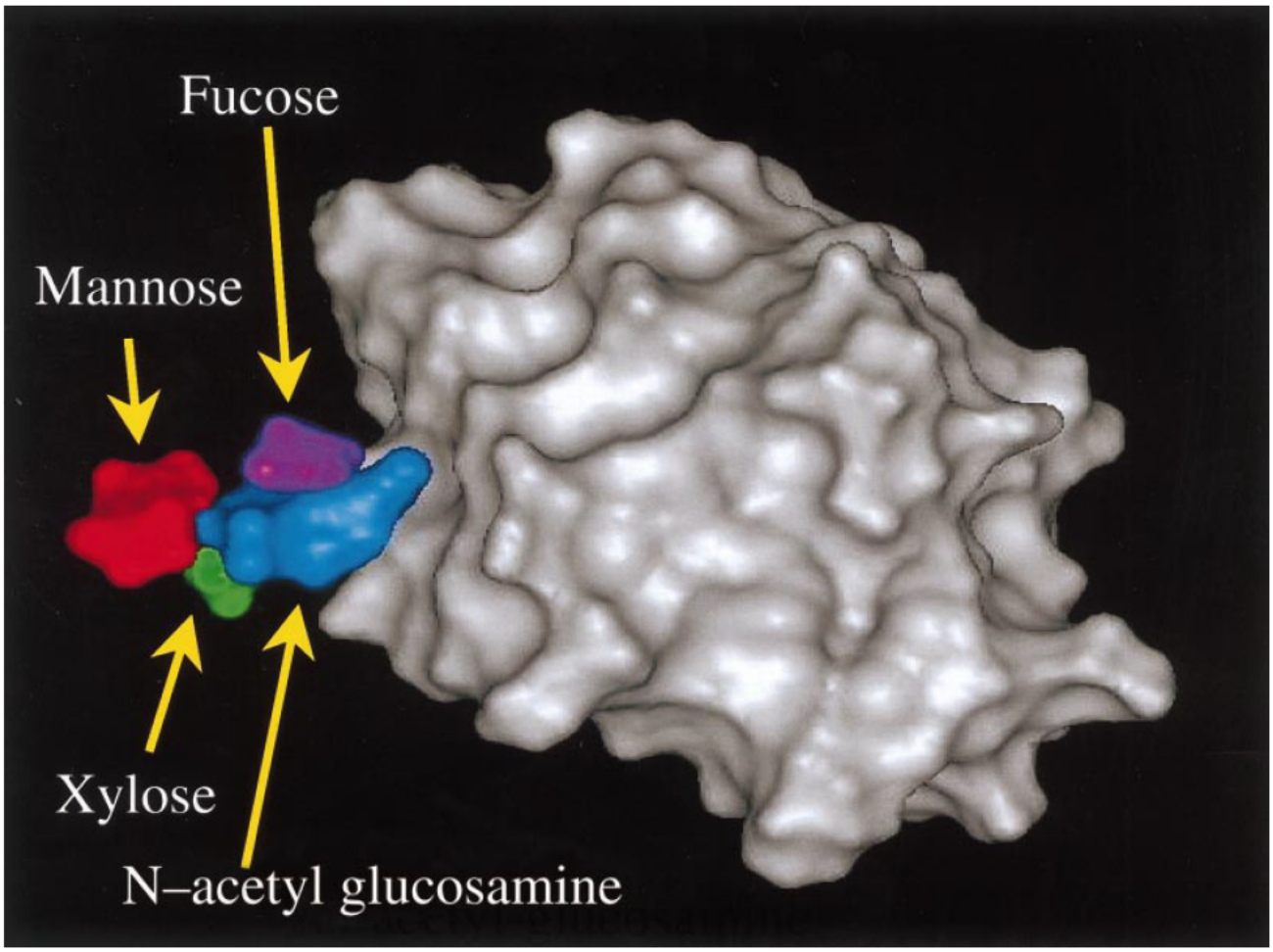


Figure 1. 3 D structure of cross reactive carbohydrate determinant (CCD) (structures in color) shown as glycan attached to its parent protein (here erythrina lectin, the gray structure) (71 72). Two sets of atomic coordinates are available from Protein Data Bank (PDB, http://www.rcsb.org/pdb/searchlite.html): PDB1LTE.ENT (shown here) and PDB2MYR.ENT, with very similar 3 D structure for glycan.

reactivity *in vitro* is not, however, always reflected by cross-reactivity *in vivo*. As mentioned before, this is particularly true for CCD-type cross-reactivity. However, such discrepancies have been found also for peptidic cross-reactivity; for example, for some patients with *in vitro* IgE reactivity to profilins, no corresponding skin reactivity is found (82, 82). This may reflect low affinity and/or restricted epitope recognition.

Immunologic sensitization in relation to cross-reactivity

Most allergic patients react to more than a single allergen. To a large degree, this reflects multiple sensitization events as a consequence of the increased immunologic responsiveness present in atopy. These events are not fully independent, as a Th2 response to one allergen is likely to facilitate subsequent Th2 responses by creating and maintaining for some time a local milieu that favors Th2 responses. In this context, the Th2-promoting potential of effector cells in the allergic cascade, particularly mast cells and basophils, is obviously of interest. This "IgE breeds IgE" (84) phenomenon and the related phenomenon of epitope spreading (85) are not as strong as sometimes suggested. It is of note that even for allergens that are simultaneously presented to the immune system, such as cat allergen and mite allergen on dust particles, cosensitization is not universal.

An alternative cause of allergic multireactivity is cross-reactivity. The definition of cross-reactivity is based on immunologic recognition. Two allergens are cross-reactive if there is a single antibody (or T-cell receptor) that reacts with both. This basic definition can be embellished by defining an affinity threshold and by including something about clinical relevance. It is important to appreciate that it may be possible to show convincingly that allergens are cross-reactive, but that it is impossible to prove that allergens are *not* cross-reactive. Allergens are non-cross-reactive until proven otherwise, and exhaustive testing is practically impossible.

In general terms, repeated contact with "the

allergen'' is a prerequisite for allergic reactions. The more accurate description is, of course, repeated contact with substances that have ''some common structural feature''. For this discussion, three points in time and three cells have to be distinguished. The first point in time is the immunologic priming. At this stage, the allergen has to be recognized by virgin T and B cells. The second point in time is the contact of the allergen with memory T and B cells. The third point in time is the contact between the allergen and the sensitized mast cell.

These three events are usually triggered by an identical allergen. How much similarity is required between the ''allergens'' on these three occasions? What happens if the allergens are only distantly related?

Cross-reactivity at the B-cell level

As discussed in more detail elsewhere (86), the surface structure recognized by antibodies is usually shaped by 6–10 amino acids, some of which are neighbors also at the primary structure level (the amino-acid sequence). A substantial contribution to the epitope usually comes from amino acids that are further apart in the linear sequence. Classically, antibodies that are reactive with isolated peptides of fewer than 15 amino acids are assumed to be directed to a linear epitope, whereas antibodies that do not react with small peptides from the linear sequence are classified as being directed to a conformational epitope. It must be stressed, however, that it is highly exceptional to find a small peptide that closely mimics the intact protein in its interaction with an antibody. There are at least two reasons why the interaction of an antibody with a linear peptide is usually several orders of magnitude weaker than with the full protein. The first is that the peptide represents only part of the epitope. The second reason is that the peptide is much more flexible than the complete protein. While such flexibility may allow the peptide to adapt its structure to an optimal fit, it decreases overall affinity because of the loss of entropy upon binding.

High-affinity cross-reactivity between proteins is thus likely to require, in general, a similar protein fold (or, rather, protein domain fold). Grafting a linear amino-acid sequence into an unrelated protein (excluding the C- or N-terminal ends) is not known to result in relevant cross-reactivity. No well-characterized example is known of high-affinity cross-reactivity between two proteins with a different fold, but with a short internal sequence of amino acids that are identical. The most probable situation where this might occur would be when the identical amino acids are either the C- or the N-terminal part in both proteins.

Many examples of single-point mutations are known that affect antibody binding drastically. This does not necessarily mean that this amino acid is the main contact residue, as a single mutation may change a much wider surface area. The interaction between antibody and antigen obviously occurs via their surfaces. This does not mean that structures below the surface are irrelevant: mutations in noncontact residues can result in a change in the configuration of contact residues and thus in a change in the affinity.

The effects described above may be quite specific for one particular antibody. On a global level, it should be obvious that if two allergens are cross-reactive when tested with one antibody, they need not be cross-reactive when tested with the next antibody. Yet, it is not always appreciated, for example, that not all IgE antibodies to profilin are cross-reactive: some react with grass profilin, and not with birch profilin (87). If a mutation affects one antibody much more than another antibody, it is tempting, but incorrect, to conclude that these two antibodies are directed to different epitopes. Similarly, if one antibody is cross-reactive and another antibody is not, these antibodies may still react with the same surface patch.

Cross-reactivity at the T-cell level

The marked difference in antigen recognition by the T-cell receptor (TCR) compared to that by antibodies obviously has implications for cross-reactivity. T-cell reactivity may be as much affected as antibody reactivity by a single amino-acid substitution. However, because of the smaller size of the peptide recognized by the T-cell receptor, the T cell is, for statistical reasons, more likely to be confronted with indistinguishable structures originating from different proteins. Moreover, because T cells are not (or, hardly) undergoing somatic mutation and affinity maturation, T-cell specificity is less adaptive than B-cell specificity. Antigen contact influences various T-cell characteristics, but not the intrinsic affinity of the TCR for the antigen. This is a major difference from antibody recognition: antibodies need to fit the antigen tightly in order to survive the second round of selection during affinity maturation.

Cross-reactivity at the mast-cell level

Both B cells and mast cells depend on antigen recognition via antibodies. Yet, cross-reactivity at the mast-cell level is distinct from that at the B-cell level (86). One important difference is the need of a B cell for additional triggers. This is in clear contrast to the more modest requirements of a mast cell. Mast-cell triggering can be modulated by various factors (adhesion molecules, cytokines, or chemokines), but these are not as essential for triggering a mast cell as

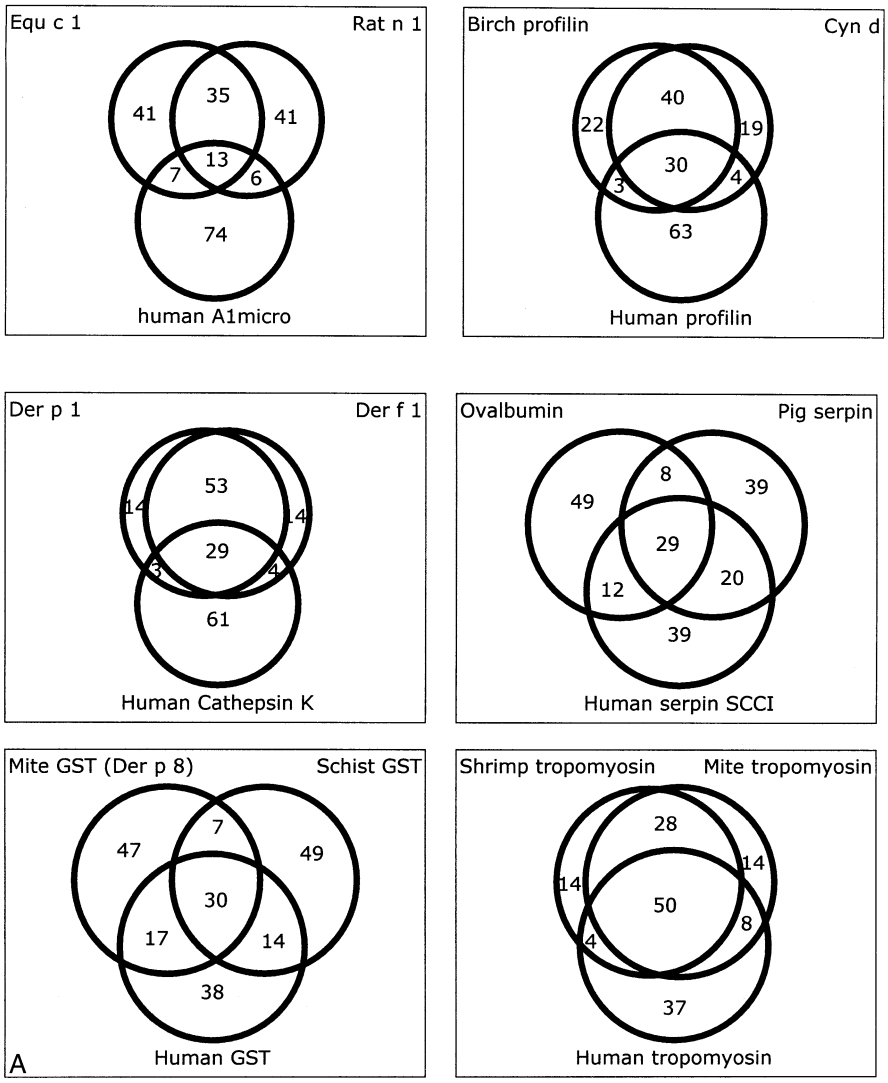


Figure 2A and B. Amino acid identity between triplets of proteins, one of which is human protein in each triplet.

they are for triggering a B cell. Moreover, the antibodies on a B-cell surface are monoclonal, whereas the antibodies on the mast-cell surface are polyclonal. The implications for cross-linking, which is the basis for activation of both B cells and mast cells, have been discussed elsewhere (86, 88).

In many cases of symptomatic cross-reactivity between allergens, it is likely that the immune system has been triggered by only one of the cross-reacting partners. The other partner reacts with cell-bound antibodies on the mast cells near the mucosal surface, but does not reach the immune system itself.

Cross-reactive priming

If the immune system has access to both cross-reactive partners, how does the first contact influence the reaction of the immune system to a subsequent contact with the cross-reactive partner? In some cases, the first contact has a tolerizing effect, as in the case of autologous proteins (Fig. 2). In this situation, part of the potential immune reactivity is downregulated or deleted; therefore, the immune response to an extraneous protein with immunologic similarity to an autologous protein is likely to be weaker than that to a completely foreign protein. If, on the other hand, the first interaction resulted in the generation of memory cells that recognize the cross-reacting antigen, an enhancing effect might be expected. This positive effect, cross-reactive priming, is more likely to occur for T cells than for B cells (see below). For T cells, this effect has been most clearly demonstrated in a classical, artificial system: carrier priming. An antihapten immune response following immunization with a hapten-carrier complex is stronger if an experimental animal has been pre-immunized with the carrier protein (89). This type of cross-reactive priming does not necessarily result in

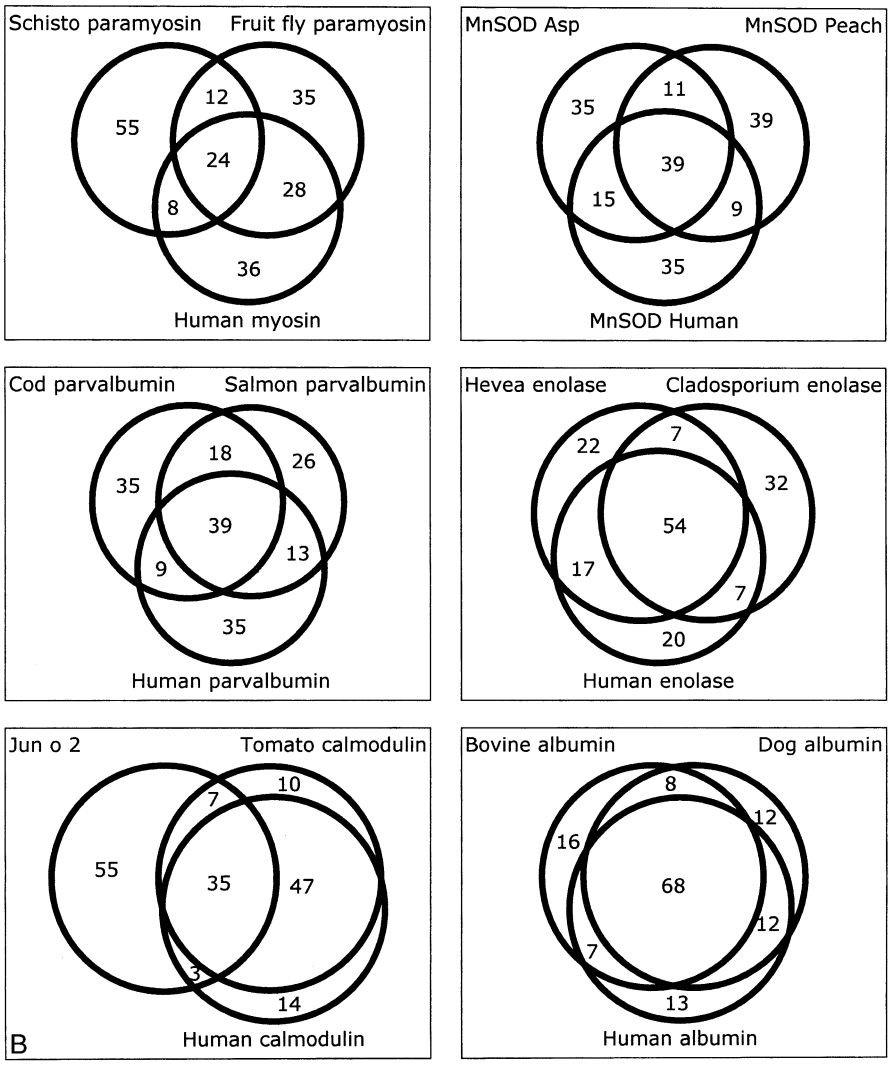


Figure 2B.

the preferential induction of cross-reacting antibodies, because the Th cells will, in principle, also stimulate B cells that are reactive with a non-cross-reactive epitope.

However, the situation is more complex in the case of B cells, because they contain a downregulatory $Fc_{gamma}RII$ receptor, which is relevant particularly in the case of naive B cells. On the one hand, just as for T cell priming, an antihapten immune response (induced by immunization with a hapten-carrier complex) is often enhanced by preimmunization with the same hapten coupled to a different carrier (90). On the other hand, the hapten (or cross-reacting epitope of the secondary antigen) will combine with the antibodies induced by the first immunization before binding to the B-cell receptor. If a naive B cell exists with a specificity for a novel (non-cross-reacting) epitope, this naive B cell will bind the immune complex. This will result in downregulation of naive B cells by the pre-existing antibodies. This effect, downregulation of an immune response by pre-existing antibodies, is the basis of the rhesus prophylaxis by administration of antirhesus antibodies.

Is the IgE response special in relation to cross-reactivity?

Cross-reactivity is certainly a prominent feature of the IgE response. Therefore, an obvious question is the following: is IgE more cross-reactive and, if so, why?

The first part of the question cannot convincingly be answered yet. There are some observations that seem to support this statement, but more data are certainly needed. Using the cat allergen Fel d 1 as a model, we found cross-reactivity with the ocelot equivalent to be more common for IgE than for IgG4 (IgE and IgG4 compared within the same serum): for IgE, the mean cross-reactivity was 66%, range 25–91%; for IgG4, it was 37%, range 14–60% ($P<0.001$) (91). Sera with IgE antibodies to *Dermatophagoides* group 1 or to grass-pollen group

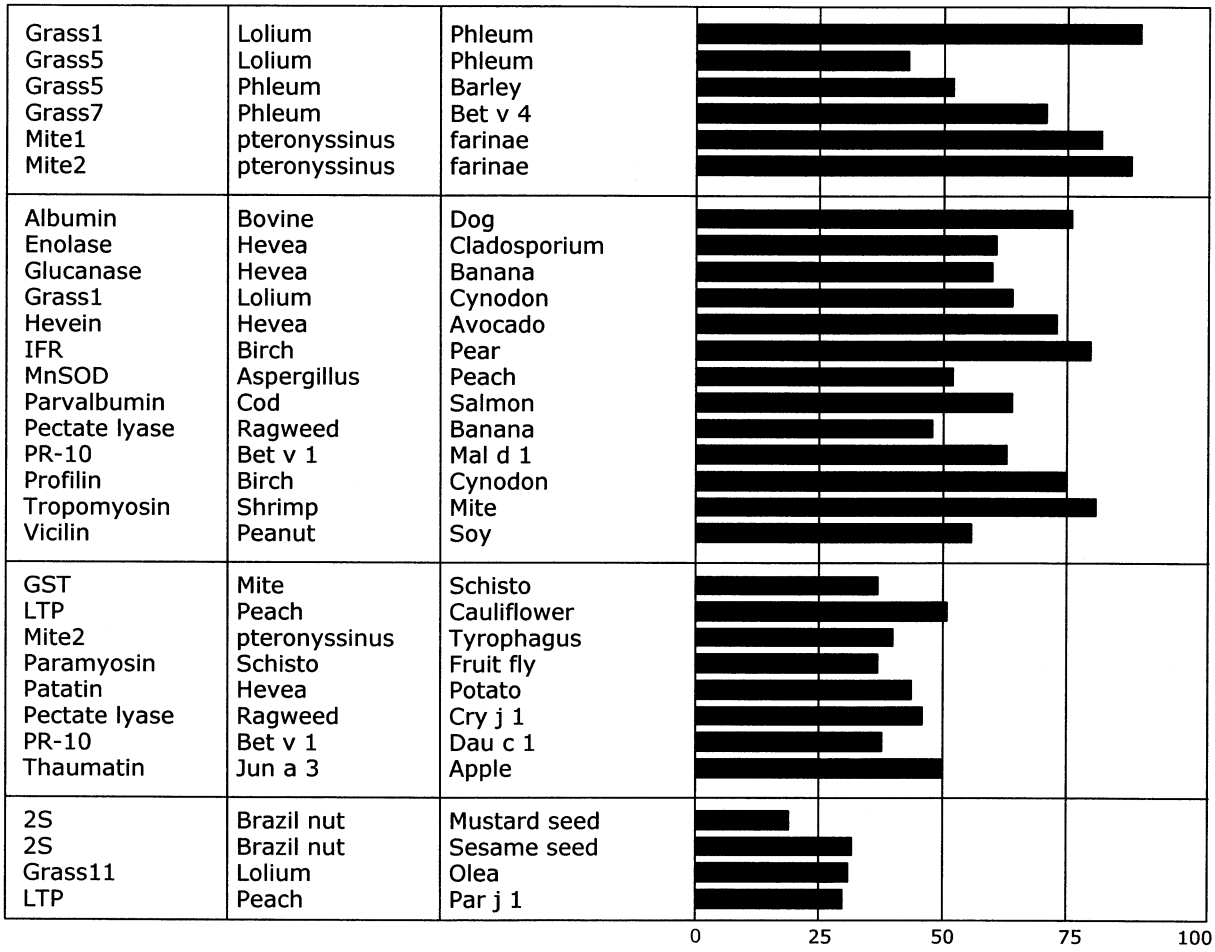


Figure 3. Amino acid identity between pairs of proteins. Proteins have been grouped into "generally cross reactive" (top panel), "often cross reactive" (middle panel), and "rarely cross reactive" (bottom panel).

1 allergens are usually cross-reactive with family members, whereas monoclonal antibodies are often found to be more specific.

If IgE is more cross-reactive (still a big "if"), why is it so? One possible answer is that this phenomenon is related to the number of antigenic contacts. In general terms, microbiologists tell us that early immune responses (disregarding the IgM response) tend to be more specific. In line with this, birch-apple cross-reactivity becomes more pronounced after long-standing pollinosis (92). This may be partially related to the phenomenon known as "the original antigenic sin" (93). This metaphor has been used to describe the lasting impression that the first influenza or HIV infection appears to have on the immune system. Upon subsequent infections with viral variants, the antibody response to the common epitopes is enhanced, whereas the response to new epitopes is suppressed, presumably due to the above-mentioned suppressive effects of pre-existing antibodies on naive B cells via $FcR_{gamma}II$. Therefore, the overall effect will be not only the expected stimulation of cross-reactive antibodies, but also suppression of non-cross-reactive antibodies. This effect, however, would explain only an increase in cross-reactivity via contacts with antigenic *variants*, whereas even repeated contact with the *same* antigen also seems to result in increased cross-reactivity. Presumably, this happens largely because of an increase in the polyclonality of the immune response, as is illustrated by the phenomenon called determinant spreading (85). The effect of polyclonality on overall cross-reactivity is a matter of statistics: the more diverse the antibody repertoire, the more likely it is to cause some cross-reactivity in an individual. If, for example, 20% of the antibodies to Bet v 1 are directed to cross-reactive epitopes and a birch-allergic patient has, on average, 10 B-cell clones activated by this allergen, the majority of patients will have a cross-reactive response, whereas most monoclonal antibodies would be non-cross-reactive. This increase in polyclonality following repeated exposure to the same antigen is a

clear indication that downregulation of naive B cells via $FcR_{gamma}II$ is not an absolute effect, since some B cells escape this downregulation. In a highly-selected model system, the human IgE response was found to be highly polyclonal (94).

In addition to the increase in polyclonality, repeated antigen contact also results in an increase in antibody affinity. Perhaps unexpectedly, this also might result in an increase in cross-reactivity. The relation between affinity and cross-reactivity is bimodal, with higher cross-reactivity among antibodies with a very low affinity as well as among those with a very high affinity, whereas antibodies with an intermediate reactivity are less cross-reactive. Low-affinity interactions often reflect a general tendency for interactions with a hydrophobic or highly charged surface patch. This type of interaction does not activate the immune regulatory circuits that result in tolerance induction. Thus, such low-affinity antibodies are often cross-reactive. As the affinity becomes higher, immune regulation becomes relevant, resulting in the downregulation of B-cell clones with reactivity to common structural motifs present in the autologous epitope repertoire. Cross-reactivity usually implies a lower affinity for the cross-reactive antigen compared to the primary antigen. Therefore, if the affinity to the primary antigen is already fairly low, the affinity for a related antigen will be even lower and thus presumably be insignificant. However, if the affinity for the primary antigen is very high, the affinity for a related antigen is much more likely to be significant.

If the formation of IgE antibodies requires repeated antigenic stimulation, the three mechanisms described above (original antigenic sin, polyclonality, and affinity maturation) may explain a relatively high prevalence of cross-reacting antibodies.

This is not an exclusive property of the IgE response. It is to be noted that the unexpected cross-reactivity pattern displayed by IgE antibodies to CCD is also not an exclusive property of IgE antibodies. Very similar cross-reactivity patterns have been described for rabbit IgG antibodies and human IgG4 antibodies (5, 95).

An interesting hypothesis has been put forward by Virtanen et al. (96) in relation to a discussion of whether proteins belonging to the lipocalin family are more likely to be allergenic. The proposal put forward was that the availability of human epitopes eliminates part of the potential immune repertoire and thus weakens the strength of the immune response. This, in turn, would favor Th2 over Th1 responses.

It is still an open question whether prior contact with an allergen via the airways increases or decreases the chance for cross-reactive responses after oral exposure to related antigens.

Can cross-reactivity be predicted?

It is tempting to try to extract rules for cross-reactivity from examples of homologous proteins with or without known IgE cross-reactivity (Fig. 3). Data on IgE cross-reactivity are, however, not always as clear-cut as one might wish. As mentioned before, complete absence of cross-reactivity cannot be proven, but theoretically a cross-reactivity scale could be made, ranging from "highly cross-reactive" (i.e., more than 50% of the IgE antibodies in a large serum pool are cross-reactive) to virtually not cross-reactive (i.e., less than 1% of the IgE antibodies are cross-reactive). For this classification to be useful, some relative affinity criterion needs to be included as well, preferably tested with soluble allergens. One approach is to use IgE binding to the cross-reactive allergen as the test system and compare the inhibition dose-response curves of the primary allergen with that of the cross-reactive allergen. If the dose required for 50% inhibition by the cross-reacting allergen were less than five times the dose required for the primary allergen, the cross-reactivity would be classified as high affinity. If the relative efficiency in this inhibition test were more than 25-fold lower, the cross-reactivity would be classified as low affinity. This is just an example of a possible approach, with a regrettable lack of scientific arguments for the factors 5 and 25. Some justification is found in the data presented for grass groups 1 and 5 (79). In this study, as in most other similar studies, no information was available on the absolute concentration of the cross-reacting allergens; therefore, differences in inhibitory efficiency could partially be explained by differences in allergen concentration.

This type of *quantitative* information is not available yet for most combinations of homologous allergens. Yet, *qualitative* information on cross-reactivity and lack of cross-reactivity is of interest as well. From the practical point of view, it is important to know whether cross-reactivity may be significant between proteins that belong to different protein-fold families. As discussed before, if two such proteins share a stretch of amino acids, this is unlikely to result in cross-reactivity unless these shared sequences both happen to be in unrestrained positions (C- or N-terminus, or perhaps a flexible loop). Within the field of allergy, no such cross-reactivity between proteins with different backbone folds has been found to result in relevant biologic effects.

Amino-acid homology is an imperfect substitute for epitope similarity. Epitopes are surface structures, whereas the most conserved part of the structure of proteins is the core. Moreover, mutations do not occur at random. Some features of a protein molecule are important for stability or function and are thus

less likely to be mutated. Yet, homology of less than 35% (less than one out of every three amino acids is identical) is unlikely to result in high-affinity cross-reactivity. High homology between two proteins is obviously more likely to result in cross-reactivity, but there is an additional factor to consider: the effect of tolerizing cross-reactivity.

The relevance of *tolerizing* cross-reactivity

The first contact with an antigen does not necessarily enhance the reaction following a subsequent contact. The most obvious factor shaping the immune repertoire is the spectrum of autologous epitopes. As a first approximation, autologous epitopes delete reactivity to epitopes on cross-reactive extraneous antigens. This tolerance is clearly not always complete, as immune reactivity to human proteins is well established (7, 97–99). Most likely, these autologous epitopes are (largely) hidden from the immune system. For the discussion of cross-reactivity patterns in general, the contribution of autologous epitopes is likely to be a strongly negative factor, i.e., a factor that limits the immunologic recognition of potentially cross-reactive structures in extraneous allergens. However, an interesting observation was made by Bernard et al. on the IgE response to β-casein (99). In 7/20 serum samples with high levels of IgE to this protein, some reactivity with the human homolog was found. This reactivity was inhibitable by the native bovine protein in solution, indicating that, in this case, the epitopes were not cryptic.

Another factor to consider is the role of tolerizing contact via the gastrointestinal tract. There is, however, not much information on food antigens suppressing sensitization to cross-reactive airborne allergens. Some potential examples exist, such as the effect of oral exposure to bovine serum albumin in milk in relation to cat (or dog) serum albumin.

Conclusions

Cross-reactivity has a clear structural basis: no relevant cross-reactivity without structural similarity. For globular proteins (i.e., for most allergens), structural similarity requires similarity in folding. Homology between two proteins that is limited to a small stretch of amino acids in the linear sequence is unlikely to result in significant cross-reactivity unless there is a similarity in the 3-D folding pattern.

Cross-reactivity means different things in different situations. Biologically, it makes a marked difference whether the cell involved is a T cell, B cell, or mast cell. More information is needed on the quantitative aspects: what is the relative avidity of the antibody toward the two antigens and what is the threshold avidity for biologic relevance of the cross-reactivity? This information is needed in order to translate immunochemically defined cross-reactivity into clinically relevant information.

Acknowledgments

We thank Erica van Oort, Ilona Kleine Budde, and Steven Stapel for allowing the use of their experimental results and for valuable discussions.

References

1. Spitzauer S. Allergy to mammalian proteins: at the borderline between foreign and self? Int Arch Allergy Immunol 1999;**120**:259 269.
2. van Ree R, Driessen MNBM, van Leeuwen WA, Stapel SO, Aalberse RC. Variability of crossreactivity of IgE antibodies to group I and V allergens in eight grass pollen species. Clin Exp Allergy 1992;**22**:611 617.
3. Anderson LBJ, Dreyfuss EM, Logan J, Johnstone DE, Glaser J. Melon and banana sensitivity coincident with ragweed pollinosis. J Allergy 1970;**45**: 310 319.
4. Hannuksela M, Lahti A. Immediate reactions to fruits and vegetables. Contact Dermatitis 1977;**3**:79 84.
5. Aalberse RC, Koshte V, Clemens JG. Immunoglobulin E antibodies that crossreact with vegetable foods, pollen, and Hymenoptera venom. J Allergy Clin Immunol 1981;**68**:356 364.
6. Valenta R, Duchene M, Ebner C, et al. Profilins constitute a novel family of functional plant pan allergens. J Exp Med 1992;**175**:377 385.
7. Valenta R, Duchene M, Pettenburger K, et al. Identification of profilin as a novel pollen allergen; IgE autoreactivity in sensitized individuals. Science 1991; **253**:557 560.
8. Kremser M, Lindemayr W. Celery allergy (celery contact urticaria syndrome) and relation to allergies to other plant antigens. Wien Klin Wochenschr 1983;**95**:838 843.
9. M'Raihi L, Charpin D, Pons A, Bongrand P, Vervloet D. Cross reactivity between latex and banana. J Allergy Clin Immunol 1991;**87**:129 130.
10. De Maat B, Akkerdaas JH, van Ree R, Aalberse RC. Vineyard snail allergy possibly induced by sensitization to house dust mite (*Dermatophagoides pteronyssinus*). Allergy 1995;**50**: 438 440.
11. van Ree R, Antonicelli L, Akkerdaas JH, Garritani MS, Aalberse RC, Bonifazi F. Possible induction of food allergy during mite immunotherapy. Allergy 1996;**51**:108 113.
12. van Ree R, Antonicelli L, Akkerdaas JH, et al. Asthma after consumption of snails in house dust mite allergic patients: a case of IgE cross reactivity. Allergy 1996;**51**:387 393.

13. Witteman AM, Akkerdaas JH, van Leeuwen J, van der Zee JS, Aalberse RC. Identification of a cross reactive allergen (presumably tropomyosin) in shrimp, mite and insects. Int Arch Allergy Immunol 1994;**105**:56 61.
14. Witteman AM, van den Oudenrijn S, van Leeuwen J, Akkerdaas J, van der Zee JS, Aalberse RC. IgE antibodies reactive with silverfish, cockroach and chironomid are frequently found in mite positive allergic patients. Int Arch Allergy Immunol 1995;**108**:165 169.
15. Björkstén F, Halmepuro L, Hannuksela M, Lahti A. Extraction and properties of apple allergens. Allergy 1980;**35**:671 677.
16. Prahl P, Weeke B, Løwenstein H. Quantitative immunoelectrophoretic analysis of extract from cow hair and dander. Characterization of the antigens and identification of the allergens. Allergy 1978;**33**:241 253.
17. Andersen KE, Løwenstein H. An investigation of the possible immunological relationship between allergen extracts from birch pollen, hazelnut, potato and apple. Contact Dermatitis 1978;**4**:73 79.
18. Calkhoven PG, Aalbers M, Koshte VL, Pos O, Oei HD, Aalberse RC. Cross reactivity among birch pollen, vegetables and fruits as detected by IgE antibodies is due to at least three distinct cross reactive structures. Allergy 1987;**42**:382 390.
19. Hsieh LS, Moos M, Lin Y. Charac terization of apple 18 and 31 kd allergens by microsequencing and evaluation of their content during storage and ripening. J Allergy Clin Immunol 1995; **96**:960 970.
20. Inschlag C, Hoffmann Sommergruber K, O'Riordain G, et al. Biochemical characterization of Pru a 2, a 23 kD thaumatin like protein representing a potential major allergen in cherry (*Prunus avium*). Int Arch Allergy Immunol 1998;**116**:22 28.
21. Karamloo F, Schmitz N, Scheurer S, et al. Molecular cloning and charac terization of a birch pollen minor allergen, Bet v 5, belonging to a family of isoflavone reductase related proteins. J Allergy Clin Immunol 1999;**104**: 991 999.
22. Pauli G, Bessot JC, Dietemann Molard A, Braun PA, Thierry R. Celery sensitivity: clinical and immuno logical correlations with pollen allergy. Clin Allergy 1985;**15**:273 279.
23. Dietschi R, Wüthrich B, Johansson SGO. Zum sogenannten "Sellerie Karotten Beifuss Gewuerz Syndrom". RAST Resultate mit neuen Gewuerzdiscs. Z Hautkr 1987;**62**: 524 531.
24. Luttkopf D, Ballmer Weber BK, Wüthrich B, Vieths S. Celery allergens in patients with positive double blind placebo controlled food challenge. J Allergy Clin Immunol 2000;**106**: 390 399.
25. Stager J, Wüthrich B, Johansson SGO. Spice allergy in celery sensitive patients. Allergy 1991;**46**:475 478.
26. Vieths S, Jankiewicz A, Wüthrich B, Baltes W. Immunoblot study of IgE binding allergens in celery roots. Ann Allergy Asthma Immunol 1995;**75**: 48 55.
27. Wüthrich B, Stager J, Johansson SGO. Celery allergy associated with birch and mugwort pollinosis. Allergy 1990;**45**: 566 571.
28. Ballmer Weber BK, Vieths S, Luttkopf D, Heuschmann P, Wüthrich B. Celery allergy confirmed by double blind, placebo controlled food challenge: a clinical study in 32 subjects with a history of adverse reactions to celery root. J Allergy Clin Immunol 2000;**106**:373 378.
29. Vallier P, Dechamp C, Vial O, Deviller P. A study of allergens in celery with cross sensitivity to mugwort and birch pollens. Clin Allergy 1988; **18**:491 500.
30. Vallier P, Dechamp C, Valenta R, Vial O, Deviller P. Purification and characterization of an allergen from celery immunochemically related to an allergen present in several other plant species. Identification as a profilin. Clin Exp Allergy 1992;**22**:774 782.
31. Jankiewicz A, Aulepp H, Baltes W, et al. Allergic sensitization to naive and heated celery root in pollen sensitive patients investigated by skin test and IgE binding. Int Arch Allergy Immunol 1996;**111**:268 278.
32. Heiss S, Fischer S, Hirschwehr R, et al. A 60 kD trypsin resistant component of mugwort pollen represents a novel cross reactive structure in pollen, fruit and vegetable allergy [Abstract]. International Congress Immunology, 1995:205 205.
33. Ganglberger E, Radauer C, Grimm R, et al. N Terminal sequences of high molecular weight allergens from celery tuber. Clin Exp Allergy 2000;**30**: 566 570.
34. Bauer L, Ebner C, Hirschwehr R, et al. IgE cross reactivity between birch pollen, mugwort pollen and celery is due to at least three distinct cross reacting allergens: immunoblot investigation of the birch mugwort celery syndrome. Clin Exp Allergy 1996;**26**:1161 1170.
35. Diaz Perales A, Lombardero M, Sanchez Monge R, et al. Lipid transfer proteins as potential plant panallergens: cross reactivity among proteins of *Artemisia* pollen, *Castanea* nut and Rosaceae fruits, with different IgE binding capacities. Clin Exp Allergy 2000;**30**:1403 1410.
36. de la Hoz F, Polo F, Moscoso DP, Selles JG, Lombardero M, Carreira J. Purification of *Art v* I, a relevant allergen of *Artemisia vulgaris* pollen. Mol Immunol 1990;**27**:651 657.
37. Breiteneder H, Ebner C. Molecular and biochemical classification of plant derived food allergens. J Allergy Clin Immunol 2000;**106**:(1 Pt 1) 27 36.
38. Sanchez Monge R, Lombardero M, Garcia Selles FJ, Barber D, Salcedo G. Lipid transfer proteins are relevant allergens in fruit allergy. J Allergy Clin Immunol 1999;**103**:514 519.
39. Pastorello EA, Farioli L, Pravettoni V, et al. The major allergen of peach (*Prunus persica*) is a lipid transfer protein. J Allergy Clin Immunol 1999; **103**:520 526.
40. Asero R, Mistrello G, Roncarolo D, et al. Lipid transfer protein: a pan allergen in plant derived foods that is highly resistant to pepsin digestion. Int Arch Allergy Immunol 2000;**122**:20 32.
41. Blanco C, Carrillo T, Castillo R, Quiralte J, Cuevas M. Latex allergy: clinical features and cross reactivity with fruits. Ann Allergy 1994;**73**: 309 314.
42. Kurup VP, Fink JN. The spectrum of immunologic sensitization in latex allergy. Allergy 2001;**56**:2 12.
43. Alenius H, Kalkkinen N, Lukka M, et al. Prohevein from the rubber tree (*Hevea brasiliensis*) is a major latex allergen. Clin Exp Allergy 1995;**25**: 659 665.
44. Sowka S, Hsieh LS, Krebitz M, et al. Identification and cloning of Prs a 1, a 32 kDa endochitinase and major allergen of avocado, and its expression in the yeast *Pichia pastoris*. J Biol Chem 1998;**273**:28091 28097.
45. Alenius H, Kalkkinen N, Reunala T, Turjanmaa K, Palosuo T. The main IgE binding epitope of a major latex allergen, prohevein, is present in its N terminal 43 amino acid fragment, hevein. J Immunol 1996;**156**:1618 1625.
46. Beezhold DH, Sussman GL, Liss GM, Chang NS. Latex allergy can induce clinical reactions to specific foods. Clin Exp Allergy 1996;**26**:416 422.
47. Breiteneder H, Sowka S, Wagner S, et al. Cloning of the patatin like latex allergen Hev b 7, its expression in the yeast *Pichia pastoris* and its immuno logical characterization. Int Arch Allergy Immunol 1999;**118**:309 310.

48. Sowka S, Wagner S, Krebitz M, et al. cDNA cloning of the 43 kDa latex allergen Hev b 7 with sequence similarity to patatins and its expression in the yeast *Pichia pastoris*. Eur J Biochem 1998;**255**:213 219.
49. Seppälä U, Alenius H, Turjanmaa K, Reunala T, Palosuo T, Kalkkinen N. Identification of patatin as a novel allergen for children with positive skin prick test responses to raw potato. J Allergy Clin Immunol 1999;**103**: 165 171.
50. Sowka S, Hafner C, Radauer C, et al. Molecular and immunologic charac terization of new isoforms of the *Hevea brasiliensis* latex allergen Hev b 7: evidence of no cross reactivity between Hev b 7 isoforms and potato patatin and proteins from avocado and banana. J Allergy Clin Immunol 1999;**104**: 1302 1310.
51. Koshte VL, Kagen SL, Aalberse RC. Crossreactivity of IgE antibodies to caddis fly with Arthropoda and Mollusca. J Allergy Clin Immunol 1989;**84**:174 183.
52. Tsai LC, Chao PL, Hung MW, et al. Protein sequence analysis and mapping of IgE and IgG epitopes of an allergenic 98 kDa *Dermatophagoides farinae* paramyosin, Der f 11. Allergy 2000; **55**:141 147.
53. van Kuik JA, van Halbeek H, Kamerling JP, Vliegenthart JFG. Primary structure of the low molecular weight carbohydrate chains of *Helix pomatia* α hemocyanin. J Biol Chem 1985;**260**:13984 13988.
54. van Kuik A, Sijbesma RP, Kamerling JP, Vliegenthart JFG, Wood EJ. Primary structure of a low molecular mass N linked oligosaccharide from hemocyanin of *Lymnaea stagnalis*. 3 *O* methyl D mannose as a constituent of the xylose containing core structure in an animal glycoprotein. Eur J Biochem 1986;**160**:621 625.
55. van Kuik JA, Hoffman RA, Mutsaers JHGM, van Halbeek H, Kamerling JP, Vliegenthart JFG. A 500 MHz 1H NMR study on the N linked carbo hydrate chain of bromelain. Glycoconj J 1986;**3**:27 34.
56. D'Andrea G, Bouwstra JB, Kamerling JP, Vliegenthart JFG. Primary structure of the xylose containing N linked carbohydrate moiety from ascorbic acid oxydase of *Cucurbita pepo medullosa*. Glycoconj J 1988;**5**:151 157.
57. Faye L, Gomord V, Fitchette Lainé AC, Chrispeels MJ. Affinity purification of antibodies specific for Asn linked glycans containing α1 3 fucose or β1 2 xylose. Anal Biochem 1993;**209**:104 108.
58. Hayashi M, Tsuru A, Mitsui T, et al. Structure and biosynthesis of the xylose containing carbohydrate moiety of rice alpha amylase. Eur J Biochem 1990;**191**:287 295.
59. Faye L, Johnson KD, Sturm A, Chrispeels MJ. Structure, biosynthesis, and function of asparagine linked glycans on plant glycoproteins. Physiologia Plantarum 1989;**75**: 309 314.
60. Ishihara H, Takahashi N, Oguri S, Tejima S. Complete structure of the carbohydrate moiety of stem bromelain. An application of the almond glycopeptidase for structural studies of glycopeptides. J Biol Chem 1979; **254**:10715 10719.
61. Kubelka V, Altmann F, Staudacher E, et al. Primary structure of the N linked carbohydrate chains from honeybee venom phospholipase A_2. Eur J Biochem 1993;**213**:1193 1204.
62. Laurière M, Laurière C, Chrispeels MJ, Johnson K, Sturm A. Charac terization of a complex xylose specific antiserum that reacts with the complex asparagine linked glycans of extra cellular and vacuolar glycoproteins. Plant Physiol 1989;**90**:1182 1188.
63. Yasuda Y, Takahashi N, Murachi T. The composition and structure of carbohydrate moiety of stem bromelain. Biochemistry 1970;**9**:25 32.
64. Lalles JP, Peltre G. Biochemical features of grain legume allergens in humans and animals. Nutr Rev 1996; **55**:101 107.
65. Staudacher E, Altmann F, März L, Hard K, Kamerling JP, Vliegenthart JFG. α1 6(α1 3) difucosylation of the asparagine bound *N* acetylglucosamine in honeybee venom phospholipase A_2. Glycoconj J 1992;**9**:82 85.
66. Takahashi N, Hotta T, Ishihara H, et al. Xylose containing common structural unit in N linked oligo saccharides of laccase from sycamore cells. Biochemistry 1986;**25**:388 395.
67. Tretter V, Altmann F, Kubelka V, März L, Becker WM. Fucose α 1,3 linked to the core region of glycoprotein *N* glycans creates an important epitope for IgE from honeybee venom allergic individuals. Int Arch Allergy Immunol 1993; **102**:259 266.
68. Fotisch K, Altmann F, Haustein D, Vieths S. Involvement of carbohydrate epitopes in the IgE response of celery allergic patients. Int Arch Allergy Immunol 1999;**120**:30 42.
69. Ueda H, Ogawa H. Glycobiology of the plant glycoprotein epitope: structure, immunogenicity and allergenicity of plant glycotopes. Trends Glycosci Glycotechnol 1999;**11**:413 428.
70. van Ree R, Cabanes Macheteau M, Akkerdaas J, et al. Beta(1, 2) xylose and alpha(1, 3) fucose residues have a strong contribution in IgE binding to plant glycoallergens. J Biol Chem 2000; **275**:11451 11458.
71. Elgavish S, Shaanan B. Structures of the *Erythrina corallodendron* lectin and of its complexes with mono and disaccharides. J Mol Biol 1998; **277**:917 932.
72. Shaanan B, Lis H, Sharon N. Structure of a legume lectin with an ordered N linked carbohydrate in complex with lactose. Science 1991;**254**:862 866.
73. van der Veen MJ, van Ree R, Aalberse RC, et al. Poor biologic activity of cross reactive IgE directed to carbo hydrate determinants of glycoproteins. J Allergy Clin Immunol 1997;**100**: 327 334.
74. Batanero E, Crespo JF, Monsalve RI, Martin Esteban M, Villalba M, Rodriguez R. IgE binding and histamine release capabilities of the main carbohydrate component isolated from the major allergen of olive tree pollen, Ole e 1. J Allergy Clin Immunol 1999;**103**:147 153.
75. Ohta M, Shigeta S, Ono K, Takao T, Shimonishi Y, Oka S. Sugar sequences of allergenically active oligosaccharide alcohols isolated from a large molecular size sea squirt antigen termed H antigen. Arch Biochem Biophys 1989;**275**:151 165.
76. Ohta M, Matsuura F, Kobayashi Y, Shigeta S, Ono K, Oka S. Further characterization of allergenically active oligosaccharitols isolated from a sea squirt H antigen. Arch Biochem Biophys 1991;**290**:474 483.
77. Shigeta S, Ono K, Ohta M, Kobayashi Y, Oka S. *N* Acetylgalactosaminyl disaccharide terminals contained in presumed carbohydrate epitopes specific to sea squirt allergy. J Biochem (Tokyo) 1989;**105**:691 696.
78. Shigeta S, Okamura M, Tsutsumi M, et al. Structure of an allergenic pentasaccharitol, Gp 1β b6, isolated from a sea squirt antigen, Gi rep, as a minimum structural unit responsible for its allergenicity. J Biochem (Tokyo) 1990;**108**:47 52.
79. van Ree R, van Leeuwen WA, Aalberse RC. How far can we simplify *in vitro* diagnostics for grass pollen allergy?: a study with 17 whole pollen extracts and purified natural and recombinant major allergens. J Allergy Clin Immunol 1998; **102**:184 190.

80. van Ree R, van Leeuwen WA, Akkerdaas JH, Aalberse RC. How far can we simplify *in vitro* diagnostics for Fagales tree pollen allergy? A study with three whole pollen extracts and purified natural and recombinant allergens. Clin Exp Allergy 1999;**29**:848–855.
81. Asero R. Effects of birch pollen specific immunotherapy on apple allergy in birch pollen hypersensitive patients. Clin Exp Allergy 1998;**28**:1368–1373.
82. Fernandez Rivas M, van Ree R, Cuevas M. Allergy to Rosaceae fruits without related pollinosis. J Allergy Clin Immunol 1997;**100**:728–733.
83. van Ree R, Fernandez Rivas M, Cuevas M, van Wijngaarden M, Aalberse RC. Pollen related allergy to peach and apple: an important role for profilin. J.Allergy Clin Immunol 1995;**95**: 726–734.
84. Aalberse RC. The IgE response and atopy. Eur Respir J 1991;**4**:78s–84s.
85. Mohapatra SS. Determinant spreading: implications in allergic disorders. Immunol Today 1994;**15**:596–597.
86. Aalberse RC. Allergen: the trigger. Allergy Clin Immunol Int 2000 **(Suppl)**:13–15.
87. van Ree R, Akkerdaas JH, van Leeuwen A, et al. New perspectives for the diagnosis of food allergy. Allergy Clin Immunol Int 2000;**12**:7–12.
88. van Ree R. Factors modulating allergen induced histamine release. In: Bousquet J, Yssel H, editors. Immunotherapy in asthma. New York: Marcel Dekker, 1999:399–409.
89. Paul WE, Katz DH, Goidl EA, Benacerraf B. Carrier function in anti hapten immune responses. II. Specific properties of carrier cells capable of enhancing anti hapten responses. J Exp Med 1970;**132**:283–299.
90. Koch G, Osmond DG, Julius MH, Benner R. The mechanism of thymus dependent antibody formation in bone marrow. J Immunol 1981;**126**: 1447–1451.
91. van Milligen FJ, van Swieten P, Aalberse RC. IgE and IgG4 binding to the ocelot variant of the cat (*Felis domesticus*) major allergen Fel d 1. Allergy 1994;**49**:393–394.
92. Asero R, Massironi F, Velati C. Detection of prognostic factors for oral allergy syndrome in patients with birch pollen hypersensitivity. J Allergy Clin Immunol 1996;**97**:611–616.
93. Fazekas de St groth B, Webster RG. Disquisitions on original antigenic sin. I. Evidence in man. J Exp Med 1966; **140**:2893–2898.
94. van't Hof W, van den Berg M, Driedijk PC, Aalberse RC. Heterogeneity in the IgE binding to a peptide derived from the house dust mite allergen *Der p* II indicates that the IgE response is highly polyclonal. Int Arch Allergy Immunol 1993;**101**:437–441.
95. van Ree R, Aalberse RC. Demonstration of carbohydrate specific IgG4 antibodies in sera of patients receiving grass pollen immunotherapy. Int Arch Allergy Immunol 1995;**106**: 146–148.
96. Virtanen T, Zeiler T, Rautiainen J, Mantyjarvi R. Allergy to lipocalins: a consequence of misguided T cell recognition of self and nonself? Immunol Today 1999;**20**:398–400.
97. Valenta R, Maurer D, Steiner R, et al. Immunoglobulin E response to human proteins in atopic patients. J Invest Dermatol 1996;**107**:203–208.
98. Mayer C, Appenzeller U, Seelbach H, et al. Humoral and cell mediated autoimmune reactions to human acidic ribosomal P2 protein in individuals sensitized to *Aspergillus fumigatus* P2 protein. J Exp Med 1999;**189**: 1507–1512.
99. Bernard H, Negroni L, Chatel JM, et al. Molecular basis of IgE cross reactivity between human beta casein and bovine beta casein, a major allergen of milk. Mol Immunol 2000;**37**:161–167.

Opinion Article

International Archives of Allergy and Immunology

Int Arch Allergy Immunol 2002;128:280–291

Received: November 12, 2001
Accepted after revision: May 14, 2002

Bioinformatic Methods for Allergenicity Assessment Using a Comprehensive Allergen Database

Ronald E. Hileman[a] Andre Silvanovich[a] Richard E. Goodman[a]
Elena A. Rice[a] Gyula Holleschak[a] James D. Astwood[a] Susan L. Hefle[b]

[a]Monsanto Company, Product Safety Center, St. Louis, Mo., and [b]University of Nebraska, Food Allergy Research and Resource Program, Lincoln, Nebr., USA

Key Words
Allergen · Allergy · *Bacillus thuringiensis* · Bioinformatics · Biotechnology · Database · Food allergy · IgE epitope · Genetically modified crops

Abstract

Background: A principal aim of the safety assessment of genetically modified crops is to prevent the introduction of known or clinically cross-reactive allergens. Current bioinformatic tools and a database of allergens and gliadins were tested for the ability to identify potential allergens by analyzing 6 *Bacillus thuringiensis* insecticidal proteins, 3 common non-allergenic food proteins and 50 randomly selected corn *(Zea mays)* proteins. ***Methods:*** Protein sequences were compared to allergens using the FASTA algorithm and by searching for matches of 6, 7 or 8 contiguous identical amino acids. ***Results:*** No significant sequence similarities or matches of 8 contiguous amino acids were found with the *B. thuringiensis* or food proteins. Surprisingly, 41 of 50 corn proteins matched at least one allergen with 6 contiguous identical amino acids. Only 7 of 50 corn proteins matched an allergen with 8 contiguous identical amino acids. When assessed for overall structural similarity to allergens, these 7 plus 2 additional corn proteins shared ≥35% identity in an overlap of ≥80 amino acids, but only 6 of the 7 were similar across the length of the protein, or shared >50% identity to an allergen. ***Conclusions:*** An evaluation of a protein by the FASTA algorithm is the most predictive of a clinically relevant cross-reactive allergen. An additional search for matches of 8 amino acids may provide an added margin of safety when assessing the potential allergenicity of a protein, but a search with a 6-amino-acid window produces many random, irrelevant matches.

Introduction

Agricultural biotechnology has been used to engineer a variety of agronomically important crops, such as corn, soybean, potato and cotton. Some are genetically modified to be resistant to insect damage. Insect-resistant varieties are generally desirable because of enhanced yield, reduced need for chemical insecticide and the high degree of pest control. Some crops that are resistant to insects have been produced by genetic modification with genes obtained from the bacterium *Bacillus thuringiensis*.

KARGER

Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1018–2438/02/1284–0280$18.50/0

Accessible online at:
www.karger.com/journals/iaa

Correspondence to: Dr. Susan L. Hefle
University of Nebraska, Food Allergy Research and Resource Program
255 Food Industry Building
Lincoln, NE 68583-0919 (USA)
Tel. +1 402 472 4430, Fax +1 402 472 1693, E-Mail shefle1@unl.edu

B. thuringiensis is a gram-positive bacterium commonly present in soil. Many different strains of *B. thuringiensis* have been shown to produce crystal proteins or inclusion bodies that are specifically effective in controlling certain orders and species of insect pests [1]. *B. thuringiensis* insecticidal proteins have been used safely for more than 40 years as biopesticides [2]. During this time, no reports of adverse human health or environmental effects have been documented [2–4]. They control larval pests, yet are safe to mammals, birds, fish and beneficial insects [3, 5]. *B. thuringiensis* crystal proteins have been generally classified based on their insecticidal activity. For example, Cry1, Cry2, Cry3 and Cry4 proteins are toxic to lepidopteran, lepidopteran/dipteran, coleopteran and dipteran pests, respectively [6, 7]. The crops that have been genetically modified to produce the *B. thuringiensis* insect control proteins Cry1Ab, Cry1Ac, Cry2Aa, Cry2Ab, Cry3Aa and Cry3Bb, are thus resistant to predation by specific insects.

Commercialization of agricultural biotechnology products requires premarket authorization. In the United States, the US Department of Agriculture, the Food and Drug Administration and the Environmental Protection Agency regulate genetically modified crops. Other countries typically have one or two agencies with similar responsibilities. Extensive food, feed and environmental safety assessments are conducted for each product prior to commercialization; this includes an allergy assessment for each introduced protein (e.g. a Cry protein) in the genetically modified crop. A principal aim is to prevent the introduction of known allergens or clinically cross-reactive incomplete allergens. An incomplete allergen is one that may elicit allergic symptoms, but does not sensitize individuals de novo [8]. The use of bioinformatic algorithms and specialized databases containing known and suspected allergens [9–13] provides important and effective tools to identify known and potentially cross-reactive allergens that may pose a risk to those with preexisting allergies. If a significant match is identified, the introduced protein must be evaluated further using sera from individuals who are allergic to the similar allergenic protein to determine if it is likely to cause cross-reactions.

Exposure to allergens in foods poses a risk to sensitized, allergic individuals. Type I (immediate hypersensitivity) reactions occur when a sufficient quantity of individual allergenic proteins, each bound by two IgE antibodies, cross-links IgE receptors on the surface of mast cells and basophils, thereby stimulating these cells to release histamine and other allergenic mediators [14]. Sensitization to produce the IgE antibodies requires a third allergen-specific recognition sequence, a CD4+ T cell epitope to provide appropriate (Th2) costimulation and cytokines to antigen specific B cells to induce IgE production [15]. Cross-reactions may be induced in sensitized individuals by a different protein if it contains IgE-binding epitopes that are identical or nearly identical to those of the sensitizing protein, as is the case for incomplete allergens, which do not include T cell epitopes for the induction of the immune response [8]. While cross-reactive structures frequently have lower affinity than the immunogen for the antibody and are therefore typically less effective at inducing an immune response [16], clinical evidence has demonstrated that foods containing proteins that share greater than 40% identity and appear to share cross-reactive IgE epitopes can induce clinical responses in some individuals [17]. However, it has not been conclusively proven that individuals with apparent cross-reactivity were not sensitized to both allergens or that those proteins are responsible for the clinical reactivity [18].

It is important to recognize that there is insufficient information to exclude the possibility that a protein that does not share sequence similarity to any known allergen will ever cause an allergic reaction. At this time there is no well-established evidence that specific allergenic motifs or structures exist that are allergenic in all contexts. Therefore, the most appropriate bioinformatic approach at this time seems to be reliance on a comparison to known allergens. Further, the additional considerations of stability to pepsin digestion and the abundance of the protein in the food may provide useful information in that regard [9, 13].

Food produced from genetically engineered crops should be as safe as that produced from traditional crops. Thus, the protection of food-allergic individuals from unexpected and unwanted exposures to offending foods or allergens is a primary health concern. Bioinformatic tools can be used to address this problem by asking two questions:

(1) Is the structure of an introduced protein sufficiently similar to a known allergen to suspect that it is an allergen or would be cross-reactive?

Current bioinformatic algorithms compare linear (local) sequences and global (three-dimensional) predictions of structure to estimate similarity. Programs such as FASTA [19] (an algorithm used to find local high-scoring alignments between a pair of protein or nucleotide sequences) and BLASTP [20] (basic local alignment search tool for proteins) can be effectively used to predict functional similarity. The basic principle of these bioinformatic algorithms is that if two proteins share sufficient linear

sequence similarity, they will also share three-dimensional structure and therefore functional homology. By definition, homologous proteins share secondary structure and common three-dimensional folds [19]. Homologous proteins are more likely to share allergenic cross-reactive conformational and linear epitopes than unrelated proteins due to similarity in sequence and structure. Because the degree of similarity between homologues varies widely, data derived from these algorithms need to be carefully evaluated in order to predict potential cross-reactivity. While related (homologous) proteins may share only 25% amino acid identity in a 200-amino-acid overlap [19], this is not generally sufficient to indicate IgE-mediated cross-reactivity [18]. Indeed, allergenic cross-reactivity caused by proteins sharing conformational or linear epitopes is rare at 50% identity and typically requires >70% amino acid identity across the full length of the protein sequences [8]. Such high levels of identity are readily detected using either FASTA or BLASTP. Additionally, proteins closely related to gliadins or glutenins, the proteins that trigger celiac disease, a non-IgE-mediated allergic disorder, can be easily identified using FASTA or BLASTP algorithms.

(2) Does the introduced protein share a short linear sequence(s) identity with an allergen?

Sequence alignment tools such as FASTA are valuable for addressing overall similarity, but proteins structurally unrelated to allergens may still contain small immunologically significant epitopes that may elicit allergic symptoms in sensitized individuals. Thus a second comparison methodology accounting for linear epitope structure has been proposed [10, 12, 13]. The presence of small (local) regions of identity could indicate the presence of common IgE-binding epitopes shared between a query sequence and a known allergen. For practical (optimal signal to noise ratio) and theoretical (approximate average linear epitope length) reasons, the current approach recommends searching for 8 or more contiguous amino acid identities [13]. In developing this approach, it was recognized that while IgE epitope maps are available for some allergens, relatively few food allergens have been fully mapped using in vitro IgE-binding studies with sera from allergic individuals. Although some linear IgE epitopes may be as short as 5 amino acids [21, 22], the majority of characterized IgE linear epitopes are 8 amino acids or longer [23]. However, only a few reports have demonstrated the avidity (of affinity) of these antibodies, or their in vivo allergic significance for these small epitopes. It is clear from some reports that high avidity binding requires 8 or more amino acids [22, 24].

All allergens do not share common structural characteristics and there is no indication of a common structure within B cell or T cell epitopes [25]. Consequently, the predictive value of a bioinformatic assessment depends on the completeness of the allergen database [11]. The identification and molecular characterization of food and other allergens have resulted in the availability of DNA and protein sequences of many important allergens in public databases and scientific literature, albeit the data are often difficult to find [11, 12]. The construction of an inclusive allergen and gliadin protein sequence database (ALLERGEN3) compiled from publicly available sequence databases and current literature is described. Further, the efficiency of searching for similarities, either by FASTA or the sliding window of 8 amino acids is greatly enhanced by using an allergen database as compared to all publicly available protein sequences.

While the focus of the bioinformatic search is on proteins that are bound by IgE, gluten-associated proteins cause celiac disease (gluten-sensitive enteropathy), a non-IgE-mediated allergic reaction, in genetically predisposed individuals. These proteins are gliadins and glutenins in wheat, triticale, rye and barley [26]. Gliadin- and glutenin-like proteins that may cause celiac disease can be identified using the same bioinformatic tools as those used to identify type I allergens. Thus clinically relevant gliadins and glutenins have been included in the allergen database.

The most recent FAO/WHO (2001) recommendations for allergenicity assessment include the use of a 6 amino acid window in a pair-wise fashion with known allergens to identify potential cross-reactive proteins in genetically modified crops used for food production [27]. In an attempt to determine if this is a justifiable alternative to the currently used 8-amino-acid window, 3 commonly consumed nonallergenic proteins and 6 proteins derived from the bacterium *B. thuringiensis* were compared to sequences in the ALLERGEN3 database using window sizes of 5–8 as well as a FASTA search. Additionally, 50 randomly selected proteins from corn *(Zea mays)* were compared to sequences in the assembled ALLERGEN3 database by scanning with a 6, 7, or 8 amino acid window and by a FASTA comparison. Corn was chosen because it is one of the least allergenic of the major grain crops and there are only one or two proteins in corn that have been identified as clinically relevant allergens [28].

Materials and Methods

All analyses were performed using the UNIX-based Genetics Computer Group (GCG) software package (version 10.0, Madison, WI) on a personal computer supported with Reflection X Client Manager network software (version 7.20, WRQ, Inc. Seattle, Wash., USA; the UNIX computer-operating system was originally termed 'UNiplexed Information and Computing Service'). Database construction was performed in the SeqLab environment of GCG. Searches were performed from an X terminal window.

Sequence Database Preparation

The allergen and gliadin sequence database (ALLERGEN3) was assembled from public domain databases including Genbank and EMBL [29], PIR [30], the NRL3D version of RCSB PBD [31, 32] and SwissProt [33]. A preliminary list of sequences was compiled from the public databases using STRINGSEARCH (keyword = allergen). Nonallergen entries were identified by reviewing published information for each entry and removed. Additional sequences were added from a previously published allergen list [13], allergens recognized by the IUIS (ftp://biobase.dk/resourcespub/who-iuis/allergen.list) and from searches of current literature available on PubMed and Entrez (http://www.ncbi.nlm.nih.gov/) [34]. For publications reporting only an N-terminal sequence, the sequence was added if 8 or more residues were provided. Duplicate sequences were identified using a FASTA search with each individual sequence. Identical entries were removed, but unique isoforms were retained (e.g. there are 35 entries of Bet v 1). The selected allergen, gliadin and glutenin sequences were compiled in the searchable database, called ALLERGEN3, containing 635 protein sequences. ALLERGEN3 is available and may be used in similarity searches at the website http://www.allergenonline.com/.

Selection of Query Protein Sequences

Six publicly available insecticidal *B. thuringiensis* protein sequences, Cry1Ab10, Cry1Ac9, Cry2Aa1, Cry2Ab2, Cry3Aa4 and Cry3Bb1 (accession No. A29125, U89872, M31738, X55416, M30503 and M89794, respectively) are representative of the *B. thuringiensis* proteins expressed in some approved genetically modified insect-resistant crops. Three proteins that are commonly consumed and not known to be allergenic were also selected for comparison (spinach rubisco large and small subunits, and bovine myoglobin). In addition, 4,116 corn protein sequences were compiled and randomly ordered from an Entrez [34] protein search using a keyword search for 'zea or maize'. The first 50 sequences found to be of 95 amino acids or greater in length were evaluated in these analyses.

Sequence Database Searches for Structural Similarity

Structural similarities shared between query sequences and the ALLERGEN3 sequence database were assessed using the FASTA algorithm [35, 36]. FASTA comparisons are initiated by aligning the first match of a specific word size (*k-tuple* parameter), in this case two. The alignment is then extended based on the chosen scoring matrix. Specific parameters used for these analyses included an expectation threshold of 10, a gap creation penalty of 12 and gap extension penalty of two. FASTA comparisons were performed using the BLOSUM50 scoring matrix [37]. Multiple alignments are made between the query sequence and each sequence in the database with a score calculated for each alignment. Only the best-scoring alignment is reported for each database sequence. The BLOSUM matrix series [37] was derived from a set of aligned, ungapped regions from protein families, called the BLOCKS database. Sequences from each block were clustered based on the percent of identical residues in the alignments [38]. The BLOSUM50 matrix will identify blocks of conserved residues that are at least 50% identical. BLOSUM50 works well for identifying sequence similarities that include gaps, and thus recognizes distant evolutionary relationships [19].

The extent of similarity was evaluated by visual inspection of the aligned sequences, the calculated percent identity and *E* score (expectation score). The *E* score reflects the degree of similarity between a pair of sequences based on matches of identical or functionally similar amino acids. Similar amino acids are structurally related and share polar, hydrophobic, or charged states. Such substitutions are referred to as 'conservative' since they are unlikely to change the structure of the protein, and by inference the function, of homologous proteins. The calculated *E* score depends on the overall length of joined (gapped) local sequence alignments, the quality (percent identity, similarity) of the overlap and the size of the database [35]. A large *E* score (e.g. >1) indicates a low degree of similarity between the query sequence and the sequence from the database. For polypeptides, a sequence similarity may indicate homology (i.e., sequences derived from a common ancestral gene with homologous function). Sequences that share extensive amino acid sequence identity and/or similarity throughout the entire alignment or identified domains will have very small *E* scores and will often share similar overall structures. Proteins that are nearly identical are very likely to share cross-reactive IgE (or IgG) and T cell epitopes.

Sequence Database Searches for Short Peptide Matches

In addition to the FASTA comparisons used to assess overall structural similarity, a pairwise comparison algorithm referred to as ALLERGENSEARCH was used to find matches of linearly contiguous amino acid identities between the query sequence and sequences within the ALLERGEN3 database. The algorithm was run from a UNIX terminal window in GCG. This program compares the query sequence to each protein sequence in a database using a sliding window of *n* amino acids, with an $n-1$ amino acid overlap relative to the preceding window. Values of '*n*' equal to 5, 6, 7 and 8 amino acids were used in this analysis, where specified.

Results

Construction of the ALLERGEN3 (Allergen and Gliadin) Database

Keyword searches of publicly available genetic databases, using the term 'allergen', were used to retrieve allergen sequences. This approach was insufficient for retrieving allergen sequences. For example, many of the entries retrieved using this method were duplicates due to the redundancy of the public databases and those were removed. Entries determined to be irrelevant based on published clinical, IgE binding or significant structural similarity data to support the declaration of allergenicity were also removed. In addition, allergen sequences were identified through a literature review and Internet searches and

Table 1. Summary results using FASTA for nonallergenic dietary proteins and selected Cry proteins

Query protein	Accession No.	Query sequence length, aa	Best-matched allergen accession No.	Allergen sequence length, aa	*E* score	Identity %	Length of overlap, aa
Bovine myoglobin	MYBO	154	Chi t 1.01 (P02229)	151	0.0028	24.8	137
Spinach rubisco SSU	P00870	123	Garden pea (Q41043)	258	0.64	42.9	21
Spinach rubisco LSU	RKSPL	475	Hev b 6 (P02877)	204	1.2	37.0	46
Cry1Ab10	A29125	1,156	Cor a 1.0102 (X71000)	160	1.0	27.6	98
Cry1Ac9	U89872	1,178	Cor a 1.0102 (X71000)	160	1.2	28.6	98
Cry2Aa1	M31738	633	Cop c 1 (CAB39376)	81	1.5	30.8	52
Cry2Ab2	X55416	633	Cop c 1 (CAB39376)	81	0.44	32.7	52
Cry3Aa4	M30503	644	Asp f 4 (CAA04959)	286	0.68	29.4	102
Cry3Bb1	M89794	652	Asp f 9 (CAA11266)	302	3.4	33.8	65

The query protein name, accession number and length is shown. Each protein was compared to the ALLERGEN3 database. The identity of the allergen found to produce the best alignment, the sequence length of that allergen (aa), expectation *(E)* score, percent identity and length of the overlapped (aligned) regions are listed.

added to the database. From this analysis, 635 protein sequence entries were deemed to be allergens. The compiled allergen sequence database is available at the web site *http://www.allergenonline.com/*, and includes links to the corresponding public database accessions. This database is considered to be conservative relative to the identification of potential allergens because putative (i.e. not clinically proven) allergens were included along with well-characterized, proven allergens. It will be important to update this database periodically as additional knowledge is gained regarding the allergenicity of the proteins currently included and as new allergens are characterized and their sequences published.

Assessment of the ALLERGEN3 Database to Identify a Known Allergen

Searching the ALLERGEN3 Database with Known Nonallergenic Sequences. Three polypeptide sequences that represent commonly consumed, nonallergenic food proteins were compared to the allergen database: the small and large subunit of spinach rubisco and bovine myoglobin. The highest-scoring similarity identified for each of these proteins is shown in table 1. The *E* score, percent identity and percent similarity values for spinach rubisco subunits did not suggest structural homology to any alleged allergen or gliadin [39]. Bovine myoglobin received a modest *E* score (0.0028) in the alignment with the *Chironomus thummi* (chironomid midge) erythrocruorin III allergen Chi t 1.01 (accession No. P02229) but shared only 24.8% amino acid identity over the 137 amino acid overlap (of 154 amino acids). These scores suggest that myoglobin and Chi t 1.01 proteins may be homologous [40]. Indeed, Chi t 1.01 is a monomeric hemoglobin and has previously been shown to share similarity to other globin proteins [41]. While this limited similarity indicates that these proteins are distantly related, the low observed percent identity does not indicate that these proteins share cross-reactive IgE-binding epitopes [18].

The sequence of each nonallergenic food protein was also compared to the ALLERGEN3 database to search for linearly contiguous amino acid identities using the ALLERGENSEARCH algorithm. None of the protein sequences contained 8 amino acid identities to any of the sequences in the ALLERGEN3 database (table 2).

Table 2. Effect of the scanning window size on the number of matches identified to allergens, using ALLERGENSEARCH

Query protein	Accession No.	Length, aa	Search window size, aa			
			5	6	7	8
Bovine myoglobin	MYBO	154	11	1[a]	0	0
Spinach rubisco SSU	P00870	123	5[b]	0	0	0
Spinach rubisco LSU	RKSPL	475	58	1[c]	0	0
Cry1Ab10	A29125	1,156	103	13	2[d]	0
Cry1Ac9	U89872	1,178	86	22	6[e]	0
Cry2Aa1	M31738	633	46	2[f]	0	0
Cry2Ab2	X55416	633	61	2[f]	0	0
Cry3Aa4	M30503	644	30	1[g]	0	0
Cry3Bb1	M89794	652	38	5[h]	0	0

The query protein names and accession numbers identify the common food protein, or bacterial insecticidal protein that were compared to the allergens in ALLERGEN3. The length of the query protein is shown in amino acids (aa). Values reported for the number of amino acids under each search window size correspond to the number of allergens that had an identical match to some segment of the query protein at the corresponding search window size.

[a] Matched allergen: bovine β-lactoglobulin (CAA32835).

[b] Matched allerens: bovine serum albumin (AAA51411), dust mite Der f 14 (BAA04558), storage mite Eur m 14 (AAF14270), midge Chi t 5 (P02224), and avocado endochitinase Per a 1 (CAB01591).

[c] Matched allergen: peanut lectin hemagglutinin (S14765).

[d] Matched allergens: garden pea pollen-like allergen (CAA59470) and cucumber expansin (AAB37746).

[e] Matched allergens: mountain cedar isoforms Jun a 1-2 (AAF80164), Jun a 1-1 (AAD03608) and Jun v 1-1 (AAF80166), garden pea pollen-like allergen (CAA59470), cucumber expansin (AAB37746) and Cup a 1 (CAB62551).

[f] Matched allergens: leucine zipper protein, mushroom Cop c 1 (CAB39376) and mustard 2S albumin Bra j 1 (P80207).

[g] Matched allergen: dust mite protein Eur m 14 (AAF14270).

[h] Matched allergens: dust mite protein Eur m 14 (AAF14270), *Aspergillus niger* xylosidase (AAD13106, CAB06417), *Aspergillus fumigatus* heat shock protein (P40292), and *Cladosporium herbarum* aldehyde dehydrogenase (S43114).

The effect of decreasing the size of the window searched below 8 contiguous amino acids on the number of hits obtained was investigated using a window size of 5, 6, 7 or 8 to search for exact matches between any portion of these query sequences and any sequence in the ALLERGEN3 database (table 2). All proteins searched using a window size of 5 amino acids matched between 5 and 58 allergens, suggesting that these hits represented noise. The number of random matches decreased dramatically when larger window sizes were used. No matches were observed for either rubisco subunit or myoglobin when a window size of 8 amino acids was used.

Assessment of B. thuringiensis *Protein Structural Similarity to Allergens and Gliadins.* Six publicly available *B. thuringiensis* protein sequences similar to those used in approved genetically modified crops were compared to the ALLERGEN3 database (table 1). Cry1Ab10 and Cry1Ac9 proteins are nearly identical (89%) to each other and both identified the *Corylus avellena* (European hazelnut) tree pollen Cor a 1.0102 (Accession No. X71000) as having the highest score. The optimized alignments for both Cry1A proteins were to the same 98-amino-acid region of Cor a 1.0102. The alignments contained multiple gaps (3 and 5, respectively), shared only 28–29% amino acid identity to Cor a 1.0102 and were short relative to the length of the Cry1A proteins (>1,150 amino acids). The low percent identity for both of these Cry1A proteins did not indicate structural homology to Cor a 1.0102 [39] as supported by the relatively large observed *E* scores (1.0 and 1.2).

The highest-scoring similarity observed to both the Cry2Aa1 and Cry2Ab2 proteins, which share 88% identity to each other, was with the *Coprinus comatus* (shaggy cap mushroom) aeroallergen Cop c 1 (accession No.

CAB39376, table 1). While approximately 31–33% amino acid identity was observed between the two Cry2 proteins and Cop c 1, the aligned region was only 52 amino acids long. These results do not indicate significant structural homology [39], again supported by the relatively large observed *E* scores (1.5 and 0.44).

The highest-scoring similarity observed to the Cry3Aa4 protein was with the *Aspergillus fumigatus* minor allergen, Asp f 4 (accession No. CAA04959, table 1). This protein has been reported to cause an allergic response only in patients with allergic bronchopulmonary aspergillosis [42]. Approximately 29% amino acid identity was shared between the Cry3Aa4 and Asp f 4 proteins in the 102-amino-acid-aligned region, with an *E* score of 0.68. The Cry3Aa4 and Cry3Bb1 protein sequences share 69% amino acid identity with each other and unlike the Cry1 and Cry2 pairs of proteins, different best-scoring allergens were identified. The highest-scoring similarity observed to Cry3Bb1 protein was with the Asp f 9 aeroallergen (accession No. CAA11266, table 1). These proteins shared approximately 34% amino acid identity with an *E* score of 3.4 for the 65 amino acid-aligned region. Neither Cry3Aa4 nor Cry3Bb1 protein was considered to be homologous to an allergenic protein due to the observed short alignments, low identity values and the relatively large observed *E* scores.

The sequence of each *B. thuringiensis* protein was also compared to the ALLERGEN3 database to search for linearly contiguous amino acid identities using the ALLERGENSEARCH algorithm (table 2). None of the 6 *B. thuringiensis* protein sequences contained 8 amino acid identities to any of the sequences in the ALLERGEN3 database. The effect of decreasing the window size below 8 contiguous amino acids on the number of hits obtained was investigated. All *B. thuringiensis* proteins searched using a window size of 5 amino acids were matched to several allergens, suggesting that these matches again represented noise. All *B. thuringiensis* proteins matched at least one allergen when a window size of 6 was used. Importantly, the number of matches with a window size of 5 increased markedly as the length of the query protein increased, suggestive of a random match phenomenon. This result was consistent with a previous report that demonstrated that matches between allergens and Cry proteins could be found at 5 and 6 contiguous identical amino acids [10]. While the Cry1A proteins matched two allergen sequences using a window of 7 amino acids, neither of these allergens was identified as having significant similarity by FASTA as the sequences of the query protein and the allergens are very dissimilar (data not shown), indicating a low probability of allergenic cross-reactivity.

Allergenicity Assessment of 50 Randomly Selected Corn (Z. mays) *Protein Sequences*

The FASTA sequence alignment tool was used to compare 50 randomly selected corn protein sequences to the ALLERGEN3 sequence database as was previously described for the *B. thuringiensis* insect control proteins. Data presented in table 2 indicated that a window size of 5 amino acids was not discriminatory. Therefore the ALLERGENSEARCH algorithm was performed using only window sizes of 6, 7 and 8 amino acids. A summary of the search results is shown in table 3.

Of the 50 sequences used to search the ALLLERGEN3 database, 10 sequences (shaded in table 3) yielded FASTA *E* scores that were likely to reflect some conservation of structure [40]. These *E* scores ranged from of 2.3×10^{-5} to 1.4×10^{-76}. Furthermore, one of these proteins displayed 100% identity along its entire length with a known corn pollen allergen [43] and could therefore be positively identified as an allergen. While the *E* score of 2.3×10^{-5} observed between zein and a wheat gliadin may suggest that these proteins share structural similarity, both are unusually rich in the amino acids glutamine and proline. This alignment was biased and included multiple inserted gaps to create the long overlap. Corn, and therefore zein, is known not to cause celiac disease. The other 8 proteins would require further analysis and testing with sera from appropriate patients to determine whether they would likely constitute a risk for individuals with specific allergies. The results of searches for identical contiguous matches of 6, 7 and 8 amino acids are also presented (table 3) and showed that 41 of the 50 query proteins contained at least 1 identical 6-amino-acid sequence match with an allergen; 15 of the 50 contained at least 1 identical 7 amino acid sequence match with an allergen. All 7 of the 50 sequences that had short linear matches of 8 amino acids with an allergen were also identified by FASTA as having significant *E* scores ($<10^{-7}$). Not all of these, however, had $\geq 35\%$ amino acid identity over ≥ 80 amino acids [27] or greater than 50% amino acids identity over the entire length of the protein [18], which have been suggested as probable lower limits for predicting cross-reactivity.

Table 3. Bioinformatic analyses of 50 randomly chosen corn proteins from public databases, comparing the results of searches for matches of 6, 7, and 8 contiguous identical amino acids (aa) with the results of a FASTA alignment search using the BLOSUM50 scoring matrix against the database of known or suspected (putative) allergens and gliadin sequences

Corn protein (Accession No.)	Corn protein (name)	Length, aa	Matches			'Best' aligned allergen using FASTA		
			6 aa	7 aa	8 aa	*E* score	% ID (aa overlap)	Accession No.
A55092[a]	catalase	493	1	0	0	0.0069	22.3 (301)	Am cockroach Per a 1, Cr-PII (U69260)
AAA33523[b]	16 kDa zein	183	11	1	0	2.3×10^{-5}	35.4 (181)	wheat gamma-gliadin B-I (M11077)
AAA68209[a]	Sus1 gene product	816	4	0	0	2.5	30.0 (60)	cherry thaumatin-like (U32440)
AAA73960[a]	kaurene synthase	823	3	0	0	2.2	30.4 (56)	hornbeam major pollen allergen (P38949)
AAA87580[a]	GAPDH	337	1	0	0	0.27	26.3 (152)	mite allergen Eur m 3 (AF047615)
AAB71078[c]	ribosomal protein P3a	120	15	7	6	5.4×10^{-8}	69.4 (49)	cladosporium 60S acidic ribosomal P2, minor allergen (P42039)
AAB86960[d]	profilin	131	20	19	19	3.6×10^{-48}	83.2 (131)	wheat profilin (X89826)
AAC72193[a]	pyruvate dehydrogenase	374	3	2	0	6.0	31.3 (48)	beta expansin
AAC78468[a]	RNA polymerase sigma factor	349	2	0	0	1.2	19.8 (96)	wheat gamma-gliadin B-III (M11335)
AAC79953[a]	anionic peroxidase H	253	2	1	0	0.096	26.7 (101)	mouse urinary protein MUP V (M16360)
AAK26754[a]	plasma membrane integral protein ZmPIP-3	292	0	0	0	0.16	26.7 (90)	apple lipid transfer protein (AAF26450)
AAK30114[a]	teosinte branched protein 1	132	2	0	0	4.3	29.2 (72)	aspergillus allergen Asp f 4 (AAF26450)
AAK51777[a]	MURB-like transposable element	207	0	0	0	1.5	28.4 (81)	mite allergen Der f 7 (Q26456)
AAK51786[a]	MURB-like protein hMURB12	207	1	0	0	2.1	28.8 (81)	mite allergen Der f 7 (Q26456)
AAK56122[d]	alpha-expansin 4	197	3	3	3	7.8×10^{-67}	81.2 (197)	pea petal protein similar to pollen allergen (X85187)
AAK56130[d]	beta-expansin 7	268	25	18	8	2.2×10^{-55}	57.0 (268)	wheat pollen allergen (U91981)
AAK59898[a]	kaurene synthase A	202	1	0	0	2.2	26.2 (65)	birch pollen allergen Bet v 1 (Z72435)
AAK60245[a]	teosinte branched protein	317	10	6	0	1.8	27.7 (83)	salmon parvalbumin (Q91483)
AAK60502[a]	sucrose export defective 1	474	2	0	0	0.14	32.8 (61)	pollen allergen 3-2 *Juniperus virginiana* (Q91483)
BAA05550[a]	Lea group 3 protein	221	1	0	0	0.11	24.9 (177)	mollusk allergen (AAB69424)
BAA22410[a]	Ca+ protein kinase-kinase	452	2	0	0	0.048	33.3 (57)	bermuda grass pollen allergen (A28050)
CAA28734[a]	40.1 kDa A1 protein	357	3	0	0	0.46	25.3 (170)	pear isoflavone reductase-related protein Pyr c 2 (AF071477)
CAA31221[a]	cytochrome oxidase su III	265	2	1	0	9.1	38.5 (26)	mold allergen Cop c 1 (AJ132235)
CAA35589[a]	pyruvate decarboxylase	610	16	4	0	0.22	22.5 (182)	cockroach Per a 1 allergen (U69957)
CAA37038[a]	dihydrodipicolinate	380	0	0	0	0.42	42.4 (33)	perennial ryegrass pollen Lol p 5 (Q40237)
CAA39438[a]	ribosomal protein S11	159	14	8	0	7.2	25.2 (111)	soy trypsin inhibitor (CAA45778)
CAA60366[a]	hypothetical plastid protein	139	0	0	0	3.7	40 (25)	para rubber tree beta glucanase (S65077)
CAA72196[a]	cytochrome p450	538	4	0	0	0.49	35.3 (51)	storage mite group 4 homologue of D pter. (AF144061)
CAA87634[a]	unknown function w/APELTALA2-like binding domain	485	5	0	0	6.4	35.1 (37)	aspergillus allergen (CAB64688)

➤

Table 3 (continued)

Corn protein (Accession No.)	Corn protein (name)	Length, aa	Matches			'Best' aligned allergen using FASTA		
			6 aa	7 aa	8 aa	*E* score	% ID (aa overlap)	Accession No.
CAB56627[a]	SBP domain protein 1	440	5	0	0	0.92	44.1 (34)	mite hsp-70 (P39674)
CAC18100[a]	putative legumain	485	1	0	0	0.045	25.9 (189)	mite allergen (P39673)
CAC35069[a]	VIP3 protein	126	0	0	0	0.05	25.9 (112)	shrimp topomyosin (U08008)
JC1524[e]	major allergen m1, pollen allergen	191	49	16	15	1.4×10^{-76}	100 (191)	corn pollen allergen (Q07154)
O24578[a]	adenylosuccinate synthetase precursor	484	2	0	0	1.6	33.3 (45)	mite group 2 (Y12690)
P04705[a]	zein alpha-precursor	186	8	0	0	0.003	28.2 (174)	wheat alpha-type gliadin (K02068)
P15590[d]	embryo globulin-1	573	7	1	0	1.1×10^{-41}	36.7 (499)	English walnut vicilin like protein (AAF18269)
P33489[a]	auxin-binding protein-5	150	11	0	0	1.4	25.8 (66)	white mustard Sin a 1 (X91799)
P40280[a]	histone H2a	159	0	0	0	0.83	30.2 (96)	fungal 60S acidic ribosomal protein (P42039)
P46251[a]	actin depolymerizing factor	139	0	0	0	0.45	20.5 (78)	canary grass pollen allergen (P56167)
P51059[a]	phosphoenolpyruvate carboxylase 2	967	0	0	0	0.52	20.9 (201)	mollusk tropomyosin
PWZMAM[a]	F1 ATPase mitochondria	508	2	0	0	1.0	25.6 (117)	fungi manganese superoxide dismutase (U53561)
S12043[a]	leucoanthocyanidin dioxygenase	395	6	0	0	0.52	26.5 (136)	canary grass pollen Pha a 5.1 (P56164)
S30062[d]	polygalacuronase	95	7	7	7	5.1×10^{-31}	70.6 (95)	polygalacturonase (CAB42866)
S37379[a]	catalase 3	496	2	0	0	0.33	20.4 (113)	wheat gamma gliadin BI (M11336)
S58532[a]	chloroplast matK protein	544	3	0	0	0.27	24 (101)	wheat alpha/beta gliadin class AII (M10092)
T02242[a]	GTP-binding protein rab2	209	2	0	0	4.2	31.6 (57)	parasitic amoeba profilin (1PRQ)
T02763[a]	probable malate dehydrogenase	652	1	0	0	0.62	23.9 (71)	canary grass pollen PHA A 5.1 (P56164)
T02990[d]	cinnamyl alcohol dehydrogenase	367	5	1	1	1.5×10^{-25}	30.4 (342)	fungi alcohol dehydrogenase (P43067)
T02993[c]	Ca+ protein kinase	531	3	0	0	4.0×10^{-11}	37.9 (145)	juniper pollen Ca+ binding protein (AF031471)
T03397[a]	hypothetical protein	132	0	0	0	0.83	23.9 (71)	mouse urinary protein V (M16360)
Totals	50		41	15	7	10 ($<10^{-4}$)	9 ($\geq$35% and $\geq$80aa overlap)	

Shaded entries correspond to those that had either $\geq$35% amino acid identities or an overlap of $\geq$80 amino acid. The matches corresponds to the number of allergens or putative allergens that matched within the specified search window. The 'Best' *E* score, % ID and allergen indicate the allergen that produced the best scoring alignment identified using the FASTA algorithm.

[a] Insignificant similarity, not a likely homologue based on the *E* score and little chance for cross-reactivity.

[b] Possible homology with the indicated allergen, low degree of identity, many gaps in alignment, little chance of cross-reactivity.

[c] Possible homology with the indicated allergen, low to moderate degree of identity, little chance of cross-reactivity.

[d] Clear homology with the indicated allergen, high degree of identity, moderate chance of cross-reactivity.

[e] Previously identified as an aeroallergen.

Discussion

Methods of biotechnology allow the insertion of genes from any source into crops that will be utilized for foods or feeds. Commercialization of these products is regulated to ensure that the foods produced from these products are as safe as foods derived from conventional crops. One important consideration is the possibility that the newly introduced gene may encode an allergen. The most significant potential risk of allergy associated with crops developed through biotechnology is the potential transfer of a known allergen from one source to another. Since proteins that are structurally very similar may be immunologically cross-reactive, it is also important to determine whether the newly introduced protein is significantly similar to any known allergen.

Comparison of the primary amino acid sequence and overall structure of the introduced protein with that of known allergens is an important step in the safety assessment process. The strategy for assessing the potential allergenicity of each protein should begin early in the developmental stages of the genetically modified plant. A decision tree describing this process has been previously reported [13]. Proteins obtained from sources with no known history of allergenicity should be evaluated using bioinformatic tools with a comprehensive allergen sequence database. Proteins that are found to be highly similar in sequence or predicted conformational structure to known allergens are then further tested with sera from individuals having allergies to the identified allergen to evaluate possible IgE binding as an indication of their allergenicity. Those proteins with specific IgE binding would not be considered further as potential products without clinical testing to prove that they are not cross-reactive.

Various computer algorithms could be used for comparing query sequences with those of known allergens. If these tools are to be predictive of allergenicity, the criteria used in the bioinformatic search must be carefully considered. The criteria for assessment of potential allergenicity of proteins produced in genetically modified crops using bioinformatic tools are currently being debated. For example, a review panel convened jointly by the FAO and WHO recommended using either a match of 6 contiguous identical amino acids, or a match of 35% identity over more than 80 amino acids to indicate that a protein is likely to be allergenic [27]. While the FASTA pairwise comparison algorithm provides a number of useful calculated values, these values are interpreted to infer homology and potential cross-reactivity. Homologous sequences share a common evolutionary history [39, 44, 45]. The percent amino acid identity is a measure of similarity for a pair of proteins. Although sequences that share a high degree of similarity (e.g. measured by percent identity) may be related, they are not necessarily homologous. In addition, a query protein that is homologous to an allergen is not itself necessarily an allergen or cross-reactive [8]. For example, bovine myoglobin (table 1), while homologous to the midge hemoglobin allergen [41], is not known to cause allergic reactions. Also, vertebrate tropomyosins, often sharing ~50% amino acid identity to known allergenic invertebrate tropomyosins (e.g. shrimp, cockroach or mite) [46], are not known to induce clinical allergic reactions.

Concerns have been raised that proteins structurally unrelated to allergens may still share an immunologically relevant epitope (a short local overlap) and may cause cross-reactions. While the number of identified allergenic epitopes is increasing, relatively few are known. A database of authenticated epitopes could be used to screen any protein for allergenic activity. However, in the absence of an extensive, empirically determined epitope database for known allergens, comparisons between overlapping peptides using algorithms such as ALLERGENSEARCH are beneficial. The ALLERGENSEARCH may be more appropriate than FASTA for identification of short potentially allergenic epitopes because it only requires that the specified window size (e.g., 8-amino-acid residues) match. While recognizing that an IgE-mediated allergic response requires at least two IgE-binding epitopes, the use of the single match of 8 identical amino acids represents a precautionary approach to identifying proteins that might cross-react and should be tested further. This search technique may be used to complement the FASTA search. FASTA comparisons are initiated by aligning a matched word size (*k-tuple* parameter) and the alignment extended based on the chosen scoring matrix. The final computed alignment is by default the highest-scoring alignment. Thus, a short local overlap may be ignored for a higher-scoring alignment containing a longer overlap but relatively few linearly contiguous amino acid identities. The use of window sizes smaller than 8 contiguous amino acids appeared to only increase the frequency of questionable positives. Together these two approaches identify similarities that may be immunologically important and relevant to the safety assessment. The use of FASTA with an inclusive allergen database will most efficiently identify known allergens and also identify proteins that are likely to share sufficient structural similarity to suspect possible cross-reactivity.

The results of the ALLERGENSEARCH (table 3) indicated that any match of 8 contiguous identical amino acids is a conservative yet appropriate criterion for such searches. The criterion of a 6 amino acid match has previously not been rigorously assessed for its predictive value and does not appear to be scientifically justified due to the high rate of matches to allergens that appear to be random. Based on these results, the FAO/WHO recommendation [27] to use a window size of 6 contiguous amino acids did not discriminate between commonly consumed nonallergenic food proteins (myoglobin and rubisco, table 2) and allergens. The use of extremely conservative criteria for the bioinformatic searches could raise unnecessary public concerns and confusion relative to the safety of our current food supply. This would also lead to significant overestimation of potential allergenicity with rejection of beneficial nonallergenic proteins or possibly necessitate considerable clinical testing to demonstrate safety. The FASTA search and additional information regarding the extent of similarity (*E* score) or the >35% identity in any segment of 80 or more amino acids [27] may provide another method that might prove more useful than strict adherence to 8 amino acid matches in the search. However, this may be overly conservative, as evidence suggests that typically greater than 50% identity is required for clinical cross-reactions [8].

None of the *B. thuringiensis* protein sequences met criteria that would indicate a need for further testing. For the 50 randomly selected corn proteins, it is interesting that the 8-amino-acid match criterion missed two allergen homologues (embryo globulin-1, P15590 and Ca+ protein kinase, T02993) and one gliadin homologue (zein, AAA33523). While each of these displayed ≥35% identity over an ≥80-amino-acid stretch, they were well below the 50% identity level, suggesting that they are unlikely to cause cross-reactions [8]. Five proteins (pyruvate dehydrogenase, AAC72193; anionic peroxidase, AAC79953; cytochrome oxidase su III, CAA31221; pyruvate decarboxylase, CAA35589 and ribosomal protein S11) matched 7 amino acids of at least one allergen that are clearly not homologues of allergens and are therefore false positives. The results of these analyses suggest that using a FASTA (or BLASTP) search, either with or without an additional 8-amino-acid match search will provide the most effective means to identify candidate product proteins that are sufficiently similar to an allergen to require further in vitro testing such as serum IgE-binding studies using sera from appropriately allergic subjects, in order to prevent the unintended introduction of allergens or cross-reactive proteins into the food supply.

As bioinformatic tools are invaluable for discovery of protein function, they are equally well suited for use in the assessment of potential allergenicity, especially in combination with other techniques [12, 13]. A comprehensive allergen database was assembled for this purpose. Overall structural similarity to known and suspected allergens should be evaluated using a method such as FASTA or BLASTP. Another sensitive screening criterion that retained selectivity for allergens or incomplete allergens versus nonallergens was a match of 8 or more contiguous identical amino acids to any known allergen. Finally, the results of these analyses demonstrate that additional work is necessary to evaluate specific threshold criteria (e.g. *E* score or between 35 and 50% identity over a significant length) for bioinformatic analysis for the prediction of allergenicity, specifically when addressing overall structural similarity and the potential for allergenic cross-reactivity.

Acknowledgments

The authors would like to thank Gary A. Bannon for critically reviewing the manuscript and Conrad H. Halling for providing the code for the algorithm used in this report and referred to as ALLERGENSEARCH. Published as Journal Series No. 13568, Agricultural Research Division, University of Nebraska-Lincoln, Lincoln, Nebr., USA.

References

1 Aronson AI, Shai Y: Why *Bacillus thuringiensis* insecticidal toxins are so effective: Unique features of their mode of action. FEMS Microbiol Lett 2001;195:1–8.
2 McClintock JT, Schaffer CR, Sjoblad RD: A comparative review of the mammalian toxicity of *Bacillus thuringiensis*-based pesticides. Pestic Sci 1995;45:95–105.
3 Siegel JP: The mammalian safety of *Bacillus thuringiensis*-based insecticides. J Invertebr Pathol 2001;77:13–21.
4 EPA: EPA registration eligibility decision (RED) *Bacillus thuringiensis.* EPA 738-R-98–004 1998;March 1998.
5 Betz FS, Hammond BG, Fuchs RL: Safety and advantages of *Bacillus thuringiensis*-protected plants to control insect pests. Regulatory Toxicol Pharmacol 2000;32:156–173.
6 de Maagd RA, Bravo A, Crickmore N: How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. Trends Genet 2001;17:193–199.

7 Bravo A: Phylogenetic relationships of *Bacillus thuringiensis* delta-endotoxin family proteins and their functional domains. J Bacteriol 1997; 179:2793–2801.
8 Aalberse RC: Structural biology of allergens. J Allergy Clin Immunol 2000;106:228–238.
9 Taylor SL, Hefle SL: Will genetically modified foods be allergenic? J Allergy Clin Immunol 2001;107:765–771.
10 Gendel SM: The use of amino acid sequence alignments to assess potential allergenicity of proteins used in genetically modified foods. Adv Food Nutr Res 1998;42:45–62.
11 Gendel SM: Sequence databases for assessing the potential allergenicity of proteins used in transgenic foods. Adv Food Nutr Res 1998;42: 63–92.
12 Astwood JD, Fuchs RL: Allergenicity of foods derived from transgenic plants. Monogr Allergy 1996;32:105–120.
13 Metcalfe DD, Astwood JD, Townsend R, Sampson HA, Taylor SL, Fuchs RL: Assessment of the allergenic potential of foods derived from genetically engineered crop plants. Crit Rev Food Sci Nutr 1996;36:S165–186.
14 Taylor SL, Hefle SL, Gauger BJ: Food allergies and sensitivities; in Helferich W, Winter CK (eds): Food Toxicology. New York, CRC Press, 2000, pp 1–36.
15 Plant M, Zimmerman EM: Allergy and mechanisms of hypersensitivity; in Paul WE (ed): Fundamental Immunology. New York, Raven Press, 1993, pp 1399–1425.
16 Berzofsky JA, Berkower IJ, Epstein SL: Antigen-antibody interactions and monoclonal antibodies; in Paul WE (ed): Fundamental Immunology. New York, Raven Press, 1993, pp 421–465.
17 Scheurer S, Wangorsch A, Nerkamp J, Skov PS, Ballmer-Weber B, Wuthrich B, Haustein D, Vieths S: Cross-reactivity within the profilin panallergen family investigated by comparison of recombinant profilins from pear (Pyr c 4), cherry (Pru av 4) and celery (Api g 4) with birch pollen profiling Bet v 2. J Chromatogr B 2001;756:315–325.
18 Aalberse RC, Akkerdaas J, van Ree R: Cross-reactivity of IgE antibodies to allergens. Allergy 2001;56:478–490.
19 Pearson WR: Flexible sequence similarity searching with the FASTA3 program package. Methods Mol Biol 2000;132:185–219.
20 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. J. Mol Biol 1990;215:403–410.
21 Beezhold DH, Hickey VL, Slater JE, Sussman GL: Human IgE-binding epitopes of the latex allergen Hev b 5. J Allergy Clin Immunol 1999; 103:1166–1172.
22 Banerjee B, Greenberger PA, Fink JN, Kurup VP: Conformational and linear B-cell epitopes of Asp f 2, a major allergen of *Aspergillus fumigatus*, bind differently to immunoglobulin E antibody in the sera of allergic bronchopulmonary aspergillosis patients. Infect Immun 1999;67:2284–2291.
23 Chatchatee P, Jarvinen KM, Bardina L, Beyer K, Sampson HA: Identification of IgE- and IgG-binding epitopes on alpha(s1)-casein: Differences in patients with persistent and transient cow's milk allergy. J. Allergy Clin Immunol 2001;107:379–383.
24 Rabjohn P, Helm EM, Stanley JS, West CM, Sampson HA, Burks AW, Bannon GA: Molecular cloning and epitope analysis of the peanut allergen Ara h 3. J Clin Invest 1999;103:535–542.
25 Bredehorst R, David K: What establishes a protein as an allergen? J Chromatogr B Biomed Sci Appl 2001;756:33–40.
26 Ferguson A: Gluten-sensitive enteropathy (celiac disease); in Metcalfe DD, Sampson HA, Simon RA (eds): Food Allergy: Adverse Reactions to Foods and Food Additives. Cambridge, Blackwell Science, 1996, pp 287–301.
27 FAO/WHO expert consultation: Evaluation of allergenicity of genetically modified foods, Rome, Jan, 2001. http://www.who.int/fsf/Gmfood/Consultation-Jan2001/report20.pdf.
28 Pastorello EA, Farioli L, Pravettoni V, Ispano M, Scibola E, Trambaioli C, Giuffrida MG, Ansaloni R, Godovac-Zimmermann J, Conti A, Fortunato D, Ortolani C: The maize major allergen, which is responsible for food-induced allergic reactions, is a lipid transfer protein. J Allergy Clin Immunol 2000;106:744–751.
29 Benson DA, Boguski MS, Lipman DJ, Ostell J: GenBank. Nucl Acids Res 1997;25:1–6.
30 George DG, Dodson RJ, Garavelli JS, Haft DH, Hunt LT, Marzec CR, Orcutt BC, Sidman KE, Srinivasarao GY, Yeh LSL, Arminski LM, Ledley RS, Tsugita A, Barker WC: The Protein Information Resource (PIR) and the PIR-International Protein Sequence Database. Nucleic Acids Res 1997;25:24–28.
31 Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: The Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
32 Bernstein FC, Koetzle TF, Williams GJ, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M: The Protein Data Bank. A computer-based archival file for macromolecular structures. Eur J Biochem 1977; 80:319–324.
33 Bairoch A, Apweiler R: The SWISS-PROT protein sequence data bank and its supplement TrEMBL. Nucleic Acids Res 1997;25:31–36.
34 Schuler GD, Epstein JA, Ohkawa H, Kans JA: Entrez: Molecular biology database and retrieval system. Methods Enzymol 1996;266:141–162.
35 Pearson WR, Lipman DJ: Improved tools for biological comparison. Proc Natl Acad Sci USA 1988;85:2440–2448.
36 Lipman DJ, Pearson WR: Rapid and sensitive protein similarity searches. Science 1985;227: 1435–1441.
37 Henikoff S, Henikoff JG: Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 1992;89:10915–10919.
38 Henikoff JG, Henikoff S: Blocks database and its applications. Methods Enzymol 1996;266: 88–105.
39 Doolittle RF: Searching through sequence databases. Methods Enzymol 1990;183:99–110.
40 Pearson WR: Effective protein sequence comparison. Methods Enzymol 1996;266:227–258.
41 Buse G, Steffens GJ, Braunitzer G, Steer W: Hemoglobins. XXV. Hemoglobin (erythrocruorin) CTT III from *Chironomus thummi thummi* (Diptera). Primary structure and relationship to other heme proteins (author's transl)]. Hoppe Seylers Z Physiol Chem 1979; 360:89–97.
42 Hemmann S, Menz G, Ismail C, Blaser K, Crameri R: Skin test reactivity to 2 recombinant *Aspergillus fumigatus* allergens in *A. fumigatus*-sentitized asthmatic subjects allows diagnostic separation of allergic bronchopulmonary aspergillosis from fungal sensitization. J Allergy Clin Immunol 1999;104:601–607.
43 Broadwater AH, Rubinstein AL, Chay CH, Klapper DG, Bedinger PA: Zea mI, the maize homolog of the allergen-encoding Lol pI gene of rye grass. Gene 1993;131:227–230.
44 Reeck GR, de Haen C, Teller DC, Doolittle RF, Fitch WM, Dickerson RE, Chambon P, McLachlan AD, Margoliash E, Jukes TH, Zuckerkandl E: 'Homology' in proteins and nucleic acids: A terminology muddle and a way out of it. Cell 1987;50:667.
45 Baxevanis, AD, Ouellette BFF (eds): Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. Methods of Biochemical Analysis.. New York, Wiley, 1998, vol 39.
46 Reese G, Ayuso R, Lehrer SB: Tropomyosin: An invertebrate pan-allergen. Int Arch Allergy Immunol 1999;119:247–258.

Research Article

Comparison of conventional FASTA identity searches with the 80 amino acid sliding window FASTA search for the elucidation of potential identities to known allergens

Gregory S. Ladics[1], Gary A. Bannon[2], Andre Silvanovich[2] and Robert F. Cressman[1]

[1] Dupont/Pioneer Crop Genetics Regulatory Science and Registration, Wilmington, DE, USA
[2] Monsanto Co., St. Louis, MO, USA

Food and Agriculture Organization/World Health Organization (FAO/WHO) recommended that IgE cross reactivity between a transgenic protein and allergen be considered when there is ≥35% identity over a sliding "window" of 80 amino acids. Our objective was to evaluate the false positive and neg ative rates observed using the FAO/WHO *versus* conventional FASTA analyses. Data used as queries against allergen databases and analyzed to assess false positive rates included: 1102 hypothetical corn ORFs; 907 randomly selected proteins; 89 randomly selected corn proteins; and 97 corn seed pro teins. To evaluate false negative rates of both methods: Bet v 1a along with several crossreacting fruit/vegetable allergens and a bean α amylase inhibitor were used as queries. Both methods were also evaluated for their ability to detect a putative nonallergenic test protein containing a sequence derived from Ara h 1. FASTA versions 3.3t0 and 3.4t25 were utilized. Data indicate a conventional FASTA analysis produced fewer false positives and equivalent false negative rates. Conventional FASTA *versus* sliding window derived *E* scores were generally more significant. Results suggest a conventional FASTA search provides more relevant identity to the query protein and better reflects the functional similarities between proteins. It is recommended that the conventional FASTA analysis be conducted to compare identities of proteins to allergens.

Keywords: Bioinformatics / FASTA analysis / Novel proteins / Protein allergen

Received: November 10, 2006; revised: February 1, 2007; accepted: March 5 2007

1 Introduction

Comparison of novel proteins for similarity to known allergens is a critical part of the weight of evidence approach used to ascertain the safety of expressed proteins in transgenic plant products. Food and Agriculture Organization/ World Health Organization (FAO/WHO) recommended in January 2001 [1] that a similarity search be performed using the FASTA algorithm [2] to search for identities in amino acid sequence that may correspond to potential IgE crossreactivity to known allergens.

When the recommendations were published, they contained a suggested procedure of how this search should be performed. One of the steps in the procedure was to "prepare a complete set of 80 amino acid length sequences derived from the expressed protein" and "compare each of the sequences" to a dataset of allergens with FASTA using a 35% or greater identity threshold over any 80 amino acid length sequences to indicate the potential for IgE crossreactivity. Based upon this suggestion, algorithms have been developed to automatically generate all possible 80 residue subpeptides from a query protein and compare each peptide against a dataset of allergens ("sliding window search"). Any 80 amino acid peptide derived from the query protein that shows ≥35% identity to a known allergen triggers the need for additional testing (*i. e.*, an IgE screening study with sera from patients allergic to the identified protein) to establish the safety of the protein in the food supply and to

Correspondence: Dr. Gregory S. Ladics, Dupont/Pioneer Crop Genetics Regulatory Science and Registration, E400/4402, Route 141 & Henry Clay Road, Wilmington, DE 19880, USA
E-mail: gregory.s.ladics@usa.dupont.com
Fax: +1-302-695-3075

Abbreviations: FAO/WHO, Food and Agriculture Organization/ World Health Organization; **FARRP,** Food Allergy Research and Resource Program; **NCBI,** National Center for Biotechnology Information

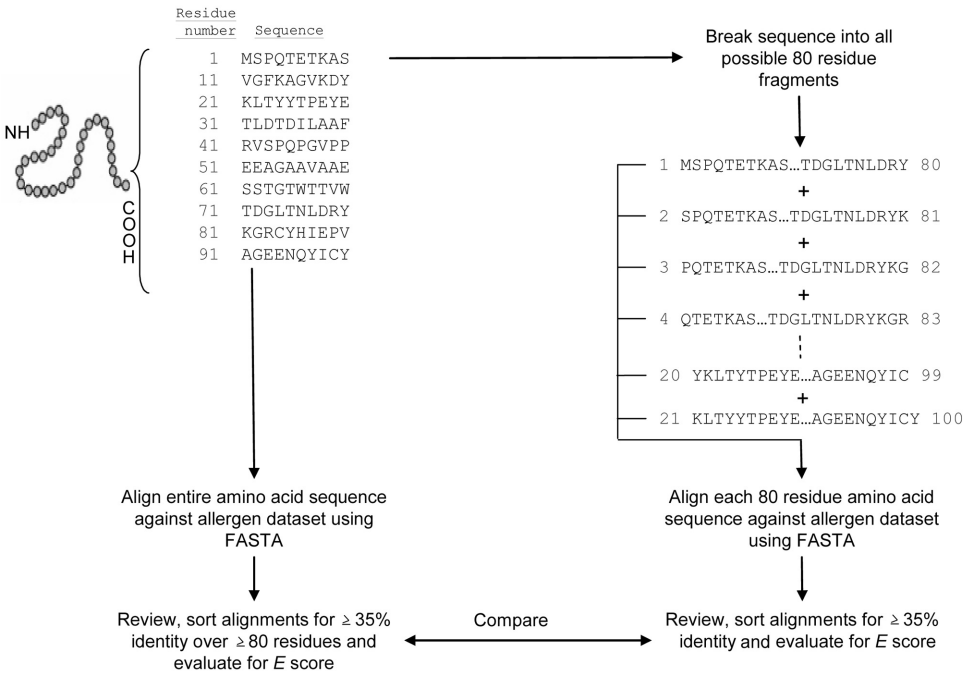


Figure 1. Schematic representation of conventional FASTA analysis of protein sequences *vs.* "sliding window" analysis. Sequences are aligned with dataset proteins conventionally using FASTA, or first broken down into all possible 80 residue "subpeptides"; each subpeptide is individually submitted for FASTA analysis. The resulting alignments are then sorted to reveal above threshold identities and *E* scores; these are then subsequently compared.

successfully register the transgenic plant product for sale in various geographies.

The use of 35% identity threshold, however, is considered to be overly conservative and likely results in a number of false positive findings. For example, others have reported that for crossreactivity to occur, a higher degree of similarity is needed, likely in excess of 50 70% sequence identity over significant spans of the target protein and allergen [3]. Radauer and Breiteneder [4] reviewed sequence identities among allergenic and nonallergenic homologs of pollen allergens and reported that the prerequisite for allergenic crossreactivity between proteins was a sequence identity of at least 50% across the length of the protein. In addition, after several years of experience in screening various proteins, we have come to believe that the use of a sliding window search in conjunction with a 35% identity threshold amplifies or exacerbates the number of potential false positive findings observed when comparing the amino acid sequences of novel proteins to those of known allergens. In order to help minimize false positive findings, a more scientifically valid approach would be to conduct the FASTA search in the conventional manner, utilizing the entire protein sequence as a query. Because the FASTA algorithm was designed to identify regions of local identity between proteins to generate an alignment, sequence along the entire length of the query protein is treated with equal weight when comparisons are made, negating the need to analyze protein sections independently. The length of sequence used to initiate and extend an aligned region is defined by the word size (k-tuple) and is set to a default value of two amino acids. Furthermore, using a sliding window search returns matches that are often inconsistent with the weight of the alignment as measured by the expectation (*E*) score. The *E* score reflects the potential random occurrence of aligned sequences and can be used to evaluate the significance of an observed alignment. The calculated *E* score depends on the overall length of gapped local sequence alignments, the quality (percent identity, similarity) of the overlap, and the size of the database [2, 5]. When comparing sequences, a very small *E* score may suggest a structurally relevant similarity, while large *E* scores (*i.e.*, >1.0) are typically associated with alignments that do not represent a biologically relevant similarity.

To test the effectiveness of these two methods, a series of *in silico* studies was performed using the Food Allergy Research and Resource Program (FARRP) Allergen Database (Version 6.0) along with a number of different protein datasets to compare the false positive and negative rates observed using a modified FASTA (*i.e.*, the 80 amino acid sliding window approach) *versus* conventional FASTA analysis at an identity threshold of 35% identity or greater (Fig. 1).

2 Methods

The following four datasets were analyzed to assess false positive rates:

(i) Hypothetical ORFs representing translations of maize genomic DNA: The ORFs were derived by using the FGENESH gene prediction program (Softberry, Mt. Kisco, NY) (data not shown). These predicted protein sequences have no homology to known proteins and there is no evidence of transcription or protein expression. For the 1102 hypothetical maize ORFs (168 sequences out of the original 1270 were less than 80 amino acids and could not generate positive hits at a length greater than 80 residues), the allergen dataset used for comparison was the Pioneer Hi-Bred International in-house allergen database. The database contained 2033 entries and was constructed by compiling protein allergen sequences identified by using keyword searches (*i.e.*, allergen(s); isoallergen(s) from published protein allergen databases [6 8] as well as the Swiss-Prot/TrEMBL, PIR, and GenPept, nr datasets). The FARRP6 Allergen Database (www.allergenonline.com; January, 2006) was employed for all other comparisons. All comparisons were carried out on an SGI MIPS R14000 computer running the IRIX Version 64 software. Briefly, either the FASTA33 or 34 programs, or a modified Perl script designed for sliding window searches (fastest33.pl, fastest34.pl) were run on all sequences in the query datasets. The sliding window scripts break a query protein sequence into all possible 80 residue subpeptides, run FASTA searches on each peptide, and return alignments that equal or exceed the FAO/WHO threshold. For all FASTA33 and 34 searches, the individual result files were concatenated into a single file; all relevant information was parsed into a summary file that was subsequently analyzed in Microsoft Excel. All matches displaying above threshold hits were identified and sorted by *E* score. Only the highest scoring match for each individual sequence was then used for the comparisons. Both conventional and sliding window FASTA searches generated additional above threshold alignments to other allergens at larger *E* scores; however, only the additional alignments generated using FASTA33 were compared in this study.

(ii) Randomly selected proteins from the Genpept dataset at National Center for Biotechnology Information (NCBI): A dataset of 1000 proteins was randomly selected from the GenePept dataset at NCBI (http://www.ncbi.nlm.nih.gov) (data not shown). Because 93 of the initial 1000 proteins were less than 80 amino acids and could not generate positive matches at a length greater than 80 residues only 907 proteins (containing both potential allergen homologs and sequences unrelated to allergens) were analyzed.

(iii) Randomly selected corn proteins: A set of amino acid sequences comprising all entries from maize (3989 accessions) were downloaded from the Uniprot dataset (http://www.pir.uniprot.org) and from this set 100 protein sequences (11 of these sequences were less than 80 residues long, the total number of sequences for percentage calculations were reduced to 89 accordingly; data not shown) were chosen at random for comparison of conventional and sliding window analysis using both FASTA33 and FASTA34.

(iv) Proteins specifically expressed in corn seed: To obtain corn protein sequences specific to seed, the edible part of the plant to which consumers are exposed, the NCBI database was searched for all proteins from corn. From this analysis, approximately 11 000 sequences were obtained. These proteins were further parsed to 248 proteins by removing hypothetical, predicted, putative, and unknown proteins, and then screening the remainder for protein sequences characterized from seed tissue. Of the 248 sequences, 14 were less than 80 amino acids while 133 were duplicates and were eliminated. This resulted in a dataset of 97 corn seed protein sequences (data not shown).

To evaluate false negatives, the following datasets were evaluated: (i) Bet v 1a and several crossreacting proteins, *i.e.*, carrot (Dau c 1); celery (Api g 1); apple (Mal d 1), cherry (Pru a 1), and pear (Pyr c 1): For the comparison of Bet v 1a and crossreacting fruit and vegetable proteins, allergen datasets were constructed that had most of the Bet v 1-like proteins removed. These datasets were then "spiked" with either the Bet v 1a protein, or the corresponding proteins from cherry (Pru a 1), celery (Api g 1), carrot (Dau c 1), apple (Mal d 1), or pear (Pyr c 1) and conventional and sliding window FASTA analysis conducted using FASTA33 and FASTA34 to determine whether key allergenic proteins would be missed using either method of FASTA analysis.

(ii) Evaluation of a bean α-amylase inhibitor transfected into pea: A nonallergenic bean α-amylase inhibitor (GI-47571317) that was transfected into pea was evaluated against the FARRP (6) Allergen Database. The bean α-amylase inhibitor expressed in the transformed pea was recently reported to display increased immunoreactivity in a nonvalidated animal model [9].

2.1 Analysis of a putative nonallergenic test protein containing target sequence from Ara h 1, a peanut allergen

To determine if, and how, a sliding window FASTA search differs from a conventional full-length FASTA search, a test protein containing a target sequence derived from Ara h 1 (GI-1168390), and a database composed of a subset of sequences derived from the FARRP (6) Allergen Database were assembled and used for comparison. The test protein sequence was composed of a single 20 amino acid segment, or a pair of variably spaced 10 amino acid segments from Ara h 1 inserted into the sequence of GI-2582631 (an acetate auxotroph from the bacteria, *Methanococcus maripaludis*). GI-2582631 was chosen due to its low degree of similarity to any sequence in the FARRP (6.0) Allergen Data-

base. The 20 amino acid target segment from Ara h 1 (amino acids 500 519) was inserted at position 60 of GI-2582631 and this sequence was used to query the FARRP Allergen Database. The 20 amino acid target segment was also split into two 10 amino acid segments (500 509 and 510 519), and these segments were inserted with variable spacing into GI-2582631 (*e.g.*, for a five amino acid spacing, one segment would be inserted at position 54 and the second at position 60) (Fig. 2). The database used for this analysis, AD6-1532, was identical to the FARRP (6.0) database with the exception that entries for Ara h 1 (GI-1168390 and -1168391) and conarachin (GI-46560472, -46560474, and -46560476) were removed. All searches were performed using FASTA version 3.3t05 on a Windows PC. Sliding window FASTA searches were implemented with FASTA Version 3.3t05 using DOS batch files.

Both older [Version 3.3t09 [FASTA33]; 2] and newer versions [Version 3.4t25 [FASTA34]; 10] of FASTA were utilized for the analysis of all datasets except for the 1270 maize ORFs and the putative nonallergenic test protein containing the target sequence from Ara h 1 (FASTA33 only).

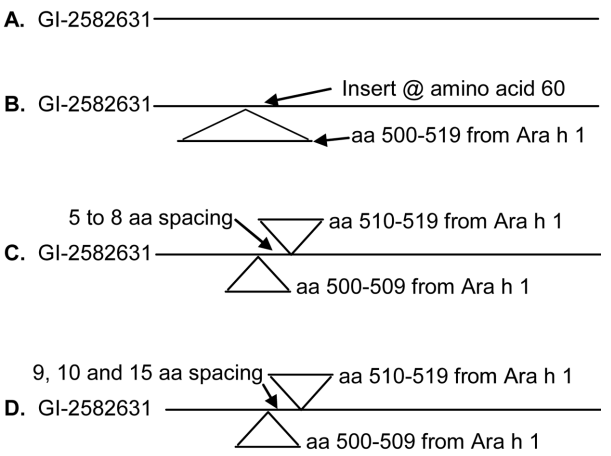


Figure 2. Analysis of a putative nonallergenic test protein containing a target sequence from the peanut allergen, Ara h 1. Construct a database consisting of AD6 minus Ara h 1 and conarachin. Conduct conventional FASTA search using A, B, and C. (A) No hits with *E* score <8. (B) Best hit is with pea vicilin. (C) Best hit with 5–8 amino acid spacing is pea vicilin. (D) Pea vicilin is no longer the top hit with conventional FASTA search. Test a sliding window search to determine if it is more sensitive than a conventional FASTA search. Seven structurally unrelated proteins including pea vicilin were identified as the top hit (depending upon the window) with a sliding window and a spacing of nine amino acids. Comparable results were obtained with 10 or 15 amino acid spacing. The sliding window search was no more reliable than the conventional FASTA search at identifying the target sequence.

3 Results

3.1 Analysis of 1102 hypothetical ORFs ≥80 amino acids from corn

A collection of ORFs (data not shown) encoding hypothetical maize proteins were subjected to allergen identity searches using an 80 amino acid sliding window FASTA33 search. A total of 73 hypothetical protein sequences out of 1102, or 6.7% of all the hypothetical protein sequences analyzed, exceeded the current threshold (*i.e.*, ≥35%) for allergenicity. When the conventional FASTA33 or FASTA34 search was used to examine the 1102 ORFs mentioned above, only 18 hypothetical protein sequences, or 1.7% of the total ORFs examined, exceeded the threshold for allergenicity (Fig. 3A). This represents approximately a five-fold decrease in the number of positive scores. These ORFs, representing translations of genomic DNA sequences using the FGENESH gene prediction program (Softberry), with no known matches to publicly available protein sequences, are unusually rich in low complexity sequences (*e.g.*, QQQQQ; PPPPPP). This could be one explanation for the dramatic difference between sliding window and conventional FASTA results when compared to the other datasets, as the sliding window search apparently magnifies the significance of low complexity sequences, as described below.

3.2 Analysis of 907 randomly selected protein sequences

When the protein sequences were subjected to the conventional FASTA analysis using either the FASTA33 or FASTA34 algorithm, 43 protein sequences (4.7%) exceeded the current recommended threshold for an identity match. In contrast, 104 protein sequences (11.5%) exceeded the threshold using a sliding window FASTA33 search, while 103 positives (11.4%) were returned when using a sliding FASTA34 search (Fig. 3B). Forty-one of the 43 sequences represented in the conventional search were present in the sliding search. *E* scores for the conventional FASTA searches were also generally much lower compared to those from the sliding window searches. For instance, the sliding window search had a larger number of *E* scores greater than 1.0 (*i.e.*, 21% of the total number of hits *vs.* 14% for the conventional FASTA33 search), which may suggest a higher number of alignments that do not represent a biologically relevant structural similarity (data not shown). These high *E* scores are reflected in the nature of the allergen hits returned using the sliding window search.

When the positive sequences from both the conventional and sliding window FASTA33 searches were analyzed for the presence of multiple above-threshold matches, approximately one half (52/104 for the sliding window search, and 19/43 for the conventional search) contained alignments to multiple FARRP sequences. The number of different allergens identified per query sequence ranged from 2 to 67, and all of the query proteins that were identified in the conventional FASTA search were contained in the sliding window

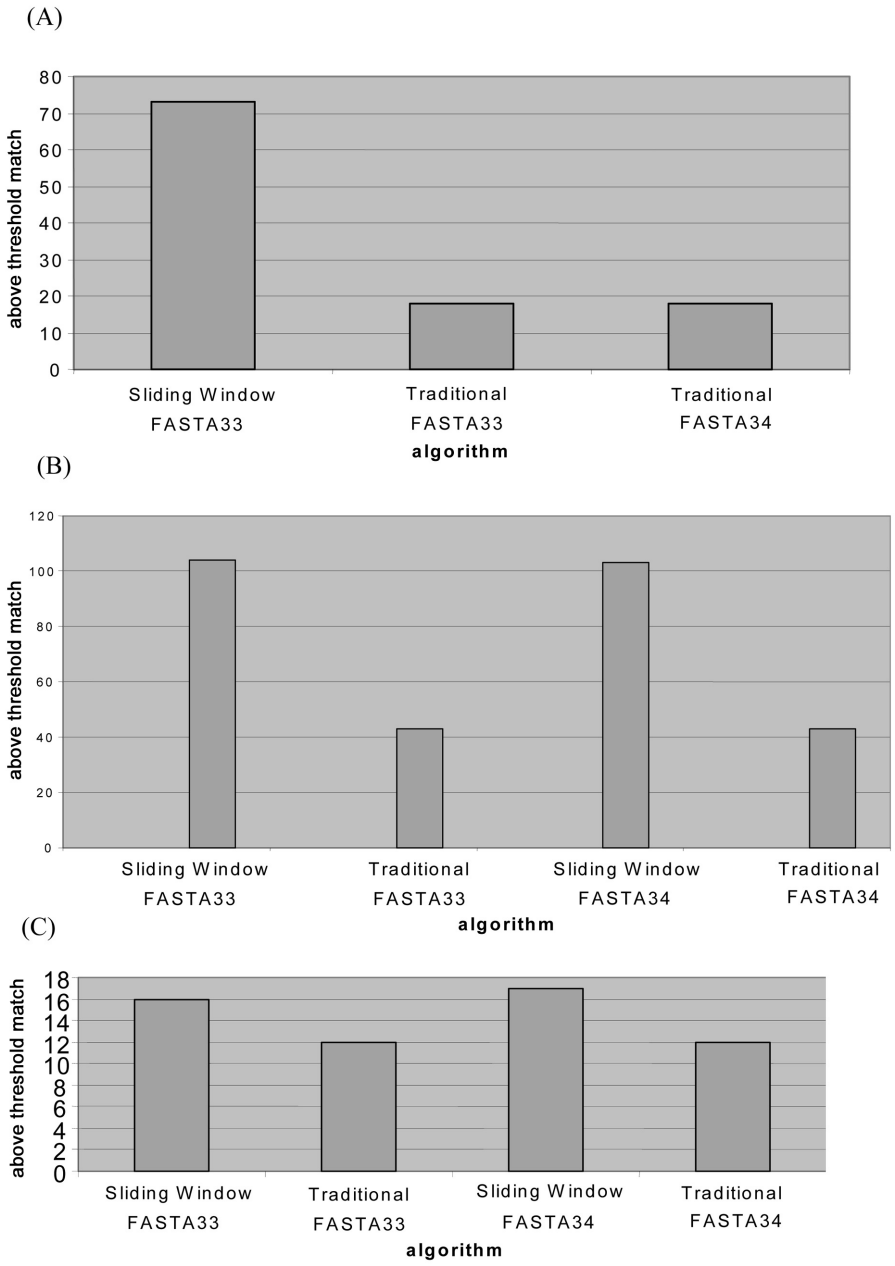


Figure 3. (A) Comparison of above threshold matches for 1102 hypothetical maize ORFs using different FASTA criteria. ORFs representing translations of maize genomic DNA encoding hypothetical maize proteins were derived by using the FGENESH gene prediction program (Softberry) and subjected to allergen identity searches using an 80 amino acid sliding window FASTA33 search or a conventional FASTA33 or FASTA34 search and the Pioneer in-house allergen database. (B) Comparison of above threshold matches for 907 randomly selected proteins using different FASTA criteria. Proteins randomly selected from the genpept dataset at NCBI were subjected to allergen identity searches using either an 80 amino acid sliding window or a conventional FASTA33 or FASTA34 search and the FARRP6 Allergen Database. (C) Comparison of above threshold matches for 89 randomly selected corn proteins using different FASTA criteria. Corn proteins randomly selected from the Uniprot dataset were subjected to allergen identity searches using either an 80 amino acid sliding window or a conventional FASTA33 or FASTA34 search and the FARRP6 Allergen Database. For all analysis, a 35% or greater identity threshold over any 80 or greater amino acid length sequence was utilized to indicate the potential for IgE crossreactivity.

Table 1. FASTA33 queries of the 907 randomly selected proteins producing multiple alignments to different allergen dataset accessions with large discrepancies between the sliding window and conventional searches

	Sliding FASTA33				Conventional FASTA33		
Query GI number	Number of matches	Avg identity (%)	Avg length (aa)	Query GI number	Number of matches	Avg identity (%)	Avg length (aa)
226743	14	36.04	85	226743	3	37.29	95
226896	12	37.20	82	226896	3	37.29	90
2267585	20	36.13	83	2267585	3	35.88	113
14268486	26	36.16	83	14268486	2	35.22	141
26985144	12	35.72	84	26985144	2	36.36	83

search. A list of proteins producing multiple alignments to different allergen dataset accessions with large discrepancies between the FASTA33 sliding window and conventional searches are listed in Table 1. For query sequences recognizing multiple allergens, the discrepency in *E* scores was even more pronounced, with only 3.1% of the conventional FASTA33 alignments producing *E* scores greater than or equal to 1, while 16% of the sliding window alignments generated to multiple allergen sequences using a single query generated *E* scores greater than or equal to 1.

One difference between the FASTA33 and FASTA34 algorithms is that the gap creation penalty has been reduced from 12 to 10, respectively. However, neither the conventional nor the sliding window results obtained using the FASTA34 algorithm displayed any significant change in the number of above threshold alignments when compared to the FASTA33 algorithm. When the conventional FASTA33 *versus* FASTA34 algorithms were compared, each produced two unique matches. The unique matches returned by the FASTA33 algorithm were close to the threshold (35.7%/84 residues and 36.4%/88 residues) and produced very high *E* scores (7.5 and 9.9, respectively). The two matches unique to the FASTA34 algorithm were also near the threshold (37.5%/80 and 35.8%/81), but returned lower *E* scores (0.0042 and 6.2, respectively). The lower *E* scores obtained with the FASTA34 algorithm suggest that the matches returned are more likely to be significant, although they would need to be further investigated. Expectation values for the FASTA34 results appear slightly lower, but in most cases the identity and length of the alignments are identical. Comparison of the FASTA34 conventional search to a sliding window search yielded results similar to those with FASTA33. Both the FASTA33 and FASTA34 sliding window searches produced 2.5 times the number of positive matches compared to the conventional FASTA searches.

3.3 Analysis of 89 randomly selected corn protein sequences

The results of the conventional *versus* sliding window analysis of the corn protein sequences using both FASTA33 and FASTA34 are summarized in Tables 2 4. An increase in the number of positives was observed with the sliding window searches compared to the conventional analysis (~19% *vs.* 12%, respectively, Fig. 3C). When extrapolated to include the total number of Uniprot derived maize accessions (*i. e.*, 3989), approximately 758 and 479 corn proteins, respectively, would be deemed potentially allergenic based on using a ≥35% identity threshold in conjunction with a sliding window and conventional FASTA analysis.

As with the 907 protein sequences mentioned above, there are several proteins that return different allergen matches, among these 89 proteins depending upon the method used; however, unlike the previous set of 907 randomly selected sequences, all allergen matches returned were similar in nature. For example, accession Q6JBQ0 ZEAMP, a chitinase, returned hits to allergenic chitinases from *Cryptomeria japonica* using the sliding window FASTA33, as well as the conventional FASTA34 search, while the sliding window FASTA34 search returned a chitinase from *Castanea sativa*; the FASTA33 conventional search returned a chitinase from *Persea americana* as the top hit. As noted previously, the conventional FASTA searches in general produced lower *E* scores (*i. e.*, <1) when compared to the sliding window searches due to the extended length of the alignments.

3.4 Analysis of 97 corn seed protein sequences

For the conventional FASTA33 or FASTA34 analyses of the corn seed proteins, 39% of these proteins were identified as potentially crossreactive with known allergenic proteins, while 53 54% were identified as allergenic using the sliding window analysis. Raising the threshold to 50% decreased the number of putative positive findings with the corn seed proteins by approximately half (17 and 25%, respectively) using either the FASTA conventional or sliding window analysis. Increasing the threshold to 70% further diminished the number of putative positive findings (3 and 6%, respectively) using either the FASTA conventional or sliding window analysis (data not shown).

Table 2. Sliding window FASTA33 above threshold allergen matches for the randomly selected corn proteins

Protein	Description	Sliding FASTA33 positives Match accession number	Allergen	Identity	Length	*E* value
GLU2 MAIZE	Glutelin-2 precursor (Zein-gamma) (27 kDa zein) (Alcohol-soluble reduced glutelin) (ASG) (Zein ZC2).	gi|170734|gb|AAA34287.1	*Gamma gliadin B-III*[b]	40.2	87	0.00017
MOSA MAIZE[a]	**Autonomous transposable element EN-1 mosaic protein (Suppressor – mutator system protein) (SPM)**	**gi|450239|gb|AAA53071.1|**	**PkIWI501**	**35**	**80**	**9.6**
O50018 MAIZE	Elongation factor 1-alpha	gi|21632054|gb|AAK85129.1|	Elongation factor (*Juniperus ashei*)	95	80	2.80E-37
Q2XXB6 ZEAMP	Pathogenesis-related protein 6	gi|62149372|dbj|BAD93486.1|	*Pollen allergen CJP38* (*C. japonica*)[b]	68.7	83	1.30E-24
Q41759 MAIZE[c]	Hypothetical protein	gi|1168391|sp|P43238|ALL12 ARAHY	Allergen Ara h 1, clone P41B precursor (Ara h 1)	35	80	0.14
Q41830 MAIZE	Mgp1 GTP-binding protein	gi|21217443|gb|AAM33785.1|	Rab11 (*Periplaneta americana*)	80	80	4.70E-33
Q41839 MAIZE	Polygalacturonase (Fragment)	gi|4826572|emb|CAB42886.1|	Polygalacturonase (*Phleum pratense*)	70	80	4.10E-30
Q41860 MAIZE	Transposable element Mu1 sequence	gi|42820661|emb|CAF31974.1|	Suppressor protein spt23-related, with ankyrin repeats (*Aspergillus fumigatus*)	35.6	90	0.75
Q4A1J1 MAIZE	cc10	gi|40807635|gb|AAR92223.1|	Phytocystatin (*Actinidia deliciosa*)	36.6	82	3.90E-05
Q64HB7 MAIZE	ASF/SF2-like pre-mRNA splicing factor SRP31	gi|63887|emb|CAA31942.1|	Vitellogenin (*Gallus gallus*)	40	80	1.50E-05
Q6JBQ0 ZEAMP	Chitinase	gi|56550550|dbj|BAD77932.1|	*Class IV chitinase* (*C. japonica*)[b]	55.4	83	7.30E-22
Q94FF5 MAIZE	Globulin 1 (Fragment)	gi|13183177|gb|AAK15089.1 |AF240006 1	7S globulin (*Sesamum indicum*)	47.5	80	3.60E-18
Q9ATL1 MAIZE	Rretrotransposon gag protein	gi|736319|emb|CAA27052.1|	Glutenin (*Triticum aestivum*)	35.8	81	0.47
Q9SWU7 MAIZE	Receptor-like kinase (Fragment)	gi|22726221|gb|AAN05083.1|	Major antigen-like protein (*Salsola kali*)	43.2	81	2.70E-11
Q9ZTL2 MAIZE	Cell wall invertase Incw1 (EC 3.2.1.26)	gi|18542113|gb|AAL75449.1 AF465612 1	Minor allergen β-fructofuranosidase precursor (*Lycopersicon esculentum*)	65	80	2.50E-26
Q9ZTQ5 MAIZE	Cell wall invertase (EC 3.2.1.26)	gi|18542113|gb|AAL75449.1| AF465612 1	Minor allergen β-fructofuranosidase precursor (*L. esculentum*)	61.3	80	4.40E-25

a) Bold text highlights differences in above threshold sequences that vary with criteria.
b) Italicized text displays allergen matches that vary with criteria used.
c) Underlined text highlights Ara h 1 similarity.

3.5 Comparison of bet v 1a and crossreacting fruit and vegetable proteins

Both the conventional and sliding window FASTA analyses correctly identified above threshold similarities between Bet v 1a and the crossreacting fruit and vegetable proteins and the false negative rates of both methods of analysis were the same (Table 5). As noted previously, the conventional FASTA resulted in lower *E* scores compared to the sliding window analysis.

Table 3. Sliding window FASTA34 above threshold allergen matches for the randomly selected corn proteins

Peptide	Description	Match accession number	Allergen	Identity	Length	*E* value
GLU2 MAIZE	Glutelin-2 precursor (Zein-gamma) (27 kDa zein) (Alcohol-soluble reduced glutelin) (ASG) (Zein ZC2).	gi|62484809|emb|-CAI78902.1	*Putative gamma-gliadin*[b] (T. aestivum)	36.1	83	0.043
O50018 MAIZE	Elongation factor 1-alpha	gi|21632054|g-b|AAK85129.1|	Elongation factor (*J. ashei*)	93.8	80	1.20E-36
Q2XXB6 ZEAMP	Pathogenesis-related protein 6	gi|1184668|g-b|AAA87456.1	*β-1,3-Glucanase*[b]	56.8	81	1.70E-15
Q41759 MAIZE[c]	Hypothetical protein	gi|1168391|sp|-P43238|ALL12 ARAHY	Allergen Ara h 1, clone P41B precursor (Ara h 1)	35	80	0.1
Q41830 MAIZE	Mgp1 GTP-binding protein	gi|21217443|g-b|AAM33785.1	Rab11 (*P. americana*)	75	80	2.40E-30
Q41839 MAIZE	Polygalacturonase (Fragment)	gi|4826572|emb|-CAB42886.1	Polygalacturonase (*P. pratense*)	70	80	1.40E-28
Q41860 MAIZE	Transposable element Mu1 sequence	gi|42820661|emb|-CAF31974.1	Suppressor protein spt23-related, with ankyrin repeats (*A. fumigatus*)	36.5	85	1.4
Q4A1J1 MAIZE	cc10	gi|40807635|g-b|AAR92223.1	Phytocystatin (*A. deliciosa*)	36.3	80	0.00012
Q64HB7 MAIZE	ASF/SF2-like pre-mRNA splicing factor SRP31	gi|63887|emb|-CAA31942.1	Vitellogenin (*G. gallus*)	40	80	8.50E-05
Q6JBQ0 ZEAMP	Chitinase	gi|1359600|emb|-CAA64868.1	*Chitinase Ib (C. sativa)*[b]	39.3	84	2.80E-08
Q94FF5 MAIZE	Globulin 1 (Fragment)	gi|13183177|gb|AA-K15089.1|AF240006 1	7S globulin (*S. indicum*)	42.7	82	3.40E-09
Q9ATL1 MAIZE	Rretrotransposon gag protein	gi|736319|emb|-CAA27052.1| glutenin (*T. aestivum*)	Glutenin (*T. aestivum*)	35.4	82	0.39
Q9ATN0 MAIZE[a]	**Plasma membrane integral protein ZmPIP1-6**	**gi|20502989|gb|AA-M22698.1|A-C098693 3**	**Putative pollen allergen (*Oryza sativa* (japonica cultivar-group))**	**35.3**	**85**	**1.9**
Q9LEE9 MAIZE	**OCL5 protein**	**gi|55859462|emb|-CAH92635.1| pollen allergen Hor v 4 H**	**Pollen allergen Hor v 4 (*Hordeum vulgare*)**	**35.3**	**85**	**4.8**
Q9SWU7 MAIZE	Receptor-like kinase (Fragment)	gi|22726221|g-b|AAN05083.1| major antigen-like protein	Major antigen-like protein (*S. kali*)	36.6	82	5.30E-08
Q9ZTL2 MAIZE	Cell wall invertase Incw1 (EC 3.2.1.26)	gi|18542113|gb|AA-L75449.1|AF465612 1 minor allergen be	Minor allergen β-fructo-furanosidase precursor (*L. esculentum*)	57.5	80	2.70E-19
Q9ZTQ5 MAIZE	Cell wall invertase (EC 3.2.1.26)	gi|18542115|gb|AA-L75450.1|AF465613 1 minor allergen be	Minor allergen β-fructo-furanosidase precursor (*L. esculentum*)	56.3	80	1.00E-19

a) Bold text highlights differences in above threshold sequences that vary with criteria.
b) Italicized text displays allergen matches that vary with criteria used.
c) Underlined text highlights Ara h 1 similarity.

3.6 Comparison of the bean α-amylase inhibitor transfected into pea to the FARP allergen database

Output from each of the two sliding window searches returned 3 above threshold identities. The highest, with identities of 35 51% over 80 residues, was to a soybean lectin (GI-170006), followed by a peanut agglutinin (GI-253289) with identities of 38 44%. Lastly, a glucose/mannose binding lectin from peanut (GI-951118) was reported, with identities between 35 and 39%. Both the conventional searches returned only the first match, with an alignment displaying 41% identity over 251 253 residues. Although not above the 35% over 80 residue threshold, the other two accessions were also captured in the conventional output.

Table 4. Randomly selected corn proteins displaying alignments to different allergen dataset accessions depending upon criteria used

Protein	FASTA33				FASTA34			
	Allergen match	% Identity	Length (aa)	*E* value	Allergen match	% Identity	Length (aa)	*E* value
Q2XXB6 ZEAMP	β-1,3-Glucanase-like protein (*Olea europaea*)	39.47	337	2.00E-16	β-1,3-Glucanase (*Hevea brasiliensis*)	53.15	333	3.8E-40
Q6JBQ0 ZEAMP	Endochitinase (*P. americana*)	44.05	311	3.20E-22	Class IV chitinase (*C. japonica*)	52.00	275	1.5E-52
Q94FF5 MAIZE	7S Globulin (*S. indicum*) cupin	36.89	225	8.50E-21	48-kDa Glycoprotein precursor (*C. avellana*) Cupin	37.17	191	1.2E-18

Though the number of hits returned varied, all four methods successfully identified the bean α-amylase inhibitor protein as a potential allergen.

3.7 Analysis of a putative nonallergenic test protein containing a target sequence from the peanut allergen, Ara h 1

When used as a query for a conventional FASTA search of AD6-1532 or FARRP (6.0) allergen databases, GI-2582631 (an acetate auxotroph from the bacteria, *M. maripaludis*) yielded no alignments with an *E* score <8 when a gap initiation penalty of 12 and gap extension penalty of 2 were employed. When a 20 amino acid target segment from Ara h 1 (amino acids 500 519 from GI-1168390) was inserted at position 60 of GI-2582631 and this sequence was used to query AD6-1532, the best alignment recovered was to *Pisum sativum* vicilin (GI-42414629). When the 20 amino acid target segment was split into two 10 amino acid segments (500 509 and 510 519), and these segments, inserted with variable spacing into GI-2582631 (for a five amino acid spacing, one segment would be inserted at position 54 and the second at position 60), a FASTA search returned *P. sativum* vicilin or a closely related homolog Len c 1.0102 (GI-29539111) as the best overall alignment for all segment spacing up to eight amino acids. Once a spacing of nine amino acids is placed between the two ten amino acid segments, a FASTA search identified a high molecular weight dust mite protein (GI-6492307) as the best alignment and *P. sativum* vicilin and Len c 1.0102 were the fifth and sixth best alignments, respectively. The conventional FASTA search was exceptionally sensitive and was able to identify the two target, ten amino acid sequences reliably with spacing of up to eight amino acids. Once the nine amino acid spacing was inserted, however, the full length FASTA search was no longer able to identify the target sequences in the tester protein.

In order to determine if a sliding window added to the sensitivity or reliability of a FASTA search, the tester protein was used as a query for 80 amino acid sliding window search. The AD6-1532 database was queried with the tester sequence that contained the ten amino acid target sequences separated by nine amino acids. The tester protein yielded a top alignment with seven structurally unrelated proteins in the AD6-1532 database depending upon the search window. These proteins included Len c 1.0102, high molecular weight dust mite protein, eosinophil granule major basic protein 2 precursor, thaumatin-like protein, ribosomal protein S12, MAG DERFA (American house dust mite allergen), and Bos d 2.0102. If each top alignment is inspected and percent identity and alignment window size are examined, the most significant of the top alignments displays 33.333% identity (36.923% ungapped) in 72 amino acid overlap with the high molecular weight dust mite protein. When 10 amino acids were used to separate the target sequences in the tester protein, the sliding 80 amino acid window FASTA search also, depending upon the window identified 7 proteins as the top alignment. However, two of the seven proteins identified relative to the test protein with the ten amino acid spacing between the target sequences differed from those having the nine amino acid spacing (*i. e.*, high molecular weight dust mite protein is no longer identified as a top alignment by any window and the top alignment displaying 24.074% identity (24.074% ungapped) in a 54 amino acid overlap is with Bos d 2.0102). Although certain windows in the sliding window search were able to identify Len c 1.0102, the use of criteria such as alignment length, or combination of length and identity to select the most significant alignment in a series of sliding window searches is no more reliable than the conventional FASTA search at identifying the target sequence. When 15 amino acids were used to separate the target sequences in the tester protein, 50 of the 88 sliding search windows identified Len c 1.0102 as the top alignment. Of the 50 alignments with Len c 1.0102, 36 alignments were the product of the insertion of a 15 amino acid gap.

4 Discussion

Comparison of the amino acid sequence of novel proteins for similarity to known or putative allergens is an important

Table 5. Comparison of Bet v 1a to crossreacting proteins using different FASTA criteria

Comparison	FASTA33 sliding window			FASTA34 sliding window			Traditional FASTA33			Traditional FASTA34		
	Identity	Length	E	Identity	Length	E	Identity	Length	E	Identity	Length	E
Betv1 *vs.* dauc1	40	80	6.50E-10	40	80	1.90E-11	38.1	155	1.00E-18	38.1	155	1.90E-20
Dauc1 *vs.* betv1	40.7	81	5.20E-11	40.7	81	4.30E-10	38.1	155	5.60E-18	38.1	155	2.10E-19
Betv1 *vs.* apig1	45	80	1.40E-10	45	80	2.10E-12	41.9	155	3.50E-21	41.9	155	3.60E-23
Apig1 *vs.* betv1	45	80	7.90E-12	45	80	2.30E-12	41.9	155	2.60E-22	41.9	155	1.70E-24
Betv1 *vs.* mald1	61.3	80	8.10E-19	61.3	80	1.60E-19	56	159	2.50E-31	56	159	2.70E-34
Mald1 *vs.* betv1	61.3	80	1.50E-21	61.3	80	6.30E-24	56	159	4.50E-28	56	159	8.50E-33
Betv1 *vs.* pyrc1	62.5	80	1.50E-18	62.5	80	3.00E-19	57.5	160	1.00E-32	57.5	160	3.70E-35
Pyrc1 *vs.* betv1	62.5	80	3.50E-22	62.5	80	6.80E-24	57.5	160	1.90E-33	57.5	160	5.10E-37
Betv1 *vs.* prua1	62.5	80	1.50E-19	62.5	80	5.30E-20	59.4	160	4.00E-35	59.4	160	1.80E-38
Prua1 *vs.* betv1	62.5	80	2.70E-21	62.5	80	1.90E-23	59.4	160	4.60E-39	59.4	160	2.50E-43

part of the safety assessment of expressed proteins in transgenic plant products. Part of this analysis involves using the FASTA algorithm [2] to search for identities in amino acid sequences that may correspond to potential IgE cross-reactivity to known or putative allergenic proteins. The objective of this study was to compare the false positive and false negative rates for two FASTA methods (*i.e.*, the sliding window *vs.* a conventional FASTA analysis). To accomplish this, a number of data sets derived from hypothetical ORFs from corn, randomly selected proteins, and corn proteins, as well as Bet v 1a homologs, an α-amylase inhibitor from bean, and a putative nonallergenic test protein containing a target sequence from the peanut allergen, Ara h 1 were utilized. Both FASTA Version 33 and 34 were employed for this comparison.

One difference between the FASTA33 and FASTA34 algorithms is that the gap creation penalty has been reduced from 12 to 10, respectively. This reduction would be expected to increase the number of gaps inserted into an alignment and therefore, increase the likelihood of any given match exceeding the FAO/WHO criteria. However, neither the conventional nor the sliding window results using the FASTA34 algorithm displayed any significant change in the number of above threshold alignments compared to FASTA33.

When a collection of ORFs encoding hypothetical maize proteins were analyzed using the sliding window search, ~7% of all sequences evaluated exceeded the current threshold of ≥35% identity, while the use of the conventional FASTA algorithm resulted in a five-fold decrease in the number of positive scores (*i.e.*, above threshold) observed with the dataset. Corn is not considered to be a major food allergen and has been classified as a "less common allergenic food" [11]. In addition, Moneret-Vautrin *et al.* [12] concluded that food allergy to corn is rare on the basis of a retrospective study on patients with histories of food allergy. The number of observed findings with the 1102 ORFs, therefore, undoubtedly reflects a large number of false positives. In addition to the five-fold increase in positive findings, the sliding window search also excluded ~13% of the sequences from analyses because they were less than 80 amino acids in length. Because the ORFs examined were hypothetical, it was not possible to determine whether any of the positive results corresponded to cross-reacting allergens. Therefore, the FASTA analysis comparison was conducted on a series of 1000 randomly selected protein sequences.

Using the conventional FASTA analysis with either the FASTA33 or FASTA34 algorithm to evaluate the 907 randomly selected proteins resulted in 2.5-fold less positive matches compared to sliding window searches. The observed percentages (4.7 and 11.5, respectively) of positive matches for the conventional and sliding window analysis, however, were higher than the expected percentage of real allergens (*e.g.*, ~0.4% for Swiss-Prot based on Swiss-Prot allergen index) [13]. This finding is likely due to the use of the currently recommended threshold of 35% [1, 14]. Data suggest that for two proteins to immunologically crossreact, a large degree of identity (in the order of 50 70%) is needed [3, 4]. *E* scores for the conventional FASTA searches were also generally much lower compared to those from the sliding window searches due to the extended

length of alignment that is possible with the conventional search. A lower *E* score may suggest a structurally relevant similarity, while large *E* scores (*e.g.*, >1.0) are typically associated with alignments that do not represent a biologically relevant structural similarity. Similar to the ORF analysis, the sliding window search excluded ~9% of the sequences from analysis because they were less than 80 amino acids in length. When the comparison between the conventional and sliding window FASTA33 analysis with the randomly selected proteins was examined further, most striking was the fact that the sliding window search resulted in 61 additional positive matches compared to the conventional analysis. Forty-one of the 43 sequences represented in the conventional search were present in the sliding search. Importantly, there were no instances identified where a sliding window search provided an informative result that differed from those obtained with a conventional search.

One of the protein sequences not present in the sliding window searches was a cytochrome oxidase from *Kradibia jacobsi*, which has similarity to a putative allergenic relative from *Sarcoptes scabiei*. This alignment generated an *E* score of 10^{-15} in the conventional searches, suggesting a high degree of potential significance. Although most proteins were represented in both the conventional and sliding window searches, in some cases the allergens matched were different. In one example, the conventional FASTA33 search identified a legume protein sequence from pea as very similar (47.5% identity over 519 residues) to an 11S globulin-like protein from *Corylus avellana*. The top match returned by the sliding window search using the same query protein sequence was to a glycinin subunit from wild soybean (*Glycine soja*). While also significant (75% identity over 80 residues), the ability to extend alignments beyond the 80 residue threshold using the conventional search generates an *E* score that is 20 orders of magnitude greater (*i.e.*, <1) than that for the sliding window search.

Although the highest scoring matches for both the sliding window and conventional searches largely matched the same accessions, analysis of the alignments produced to multiple distinct allergen dataset acccessions revealed five separate instances where a large number of positives were returned when compared to the conventional FASTA33 search (Table 1). A closer examination of the individual alignments reveals the additional matches generated by the sliding window search are based upon multiple stretches of low complexity sequence, such as QQQQ, PPPP, or EEEE. Within the context of a conventional search, these short sequences are part of a larger alignment window (the range for the examples is from 83 to 141 amino acids). In contrast, when the window size is reduced to 80 residues (range is from 82 to 85 amino acids for the sliding search), these regions are of greater influence, resulting in an increased number of above threshold alignments to sequences containing short stretches of matching sequence. These short repetive sequences are the hallmark of many celiac proteins, such as the gliadin and glutenins from wheat, which make up the majority of the additional positives returned. Sliding window positive alignments can also be generated based upon short matching "words" that are not repetitive. An example of this is found when a heat shock protein from *Bradyrhizobium* (GI 12642164) is used as a query protein. Both the sliding window and conventional searches return multiple hits including the allergens Cla h 4 and Pen c 19, both of which are heat shock proteins, but the sliding window search returns an additional alignment to a tropomyosin from cockroach. This single alignment has a much higher *E* score than the others (4.2 *vs.* 10^{-26} for the heat shock proteins) and appears to be due to the presence of the matching peptide AEADKK at the beginning of the alignment.

Based on these data, it appears that subjecting a protein to a sliding window search is more likely to result in a match that, in some cases, is not functionally related to the query protein. Crossreactive allergens are typically functionally/structurally related. For example, Breiteneder and Ebner [15] reported that plant food allergens are either homologous to pathogenesis-related-type proteins or belong to a small number of protein classes, such as seed storage proteins or enzyme inhibitors. Mills *et al.* [16] further indicated that plant food allergens are members of three structurally related superfamilies that include: the prolamin superfamily (2S albumins, nonspecific lipid transfer proteins, and cereal α-amylase/trypsin inhibitors), the cupin superfamily (7S and 11S storage proteins from peanut, soy and tree nuts), and cysteine proteases (papain-like proteases). The majority of plant food allergens are either protective or storage proteins [17]. Jenkins *et al.* [18] further confirmed these findings by determining that the majority of plant food allergens belong to only four structural families (*i.e.*, prolamin, Betv 1 family, cupin, and profilins) accounting for over 65% of food allergens. Bredehors and David [19] also concluded that functional aspects of some allergens might play a role in the allergic response. To date, IgE crossreactivity between structurally unrelated allergens has not been demonstrated [20].

In order to compare the conventional and sliding window methodologies with a food not considered to be a major allergen [11], 89 random sequences from corn and 97 sequences specific to corn seed (*i.e.*, the edible part to which consumers are exposed) were obtained. Extrapolating the data from the 89 random sequences obtained from corn indicated that approximately 758 (19%) and 479 (12%) of Uniprot derived maize accessions, respectively would be identified as potentially allergenic based on the sliding window and conventional FASTA analysis. These data are similar to the percentage (*i.e.*, 18%) observed by Hileman *et al.* [21] in which 50 randomly selected corn proteins were evaluated using the conventional FASTA analysis. Similarly, 39% of corn seed proteins were identified as

```
>>gi|1168391|sp|P43238|ALL12_ARAHY Allergen Ara h 1, clo  (626 aa)
 initn:  75 init1:  50 opt:  88  Z-score: 117.0  bits: 28.8 E(): 0.16
Smith-Waterman score: 88; 35.000% identity (43.077% ungapped) in 80 aa
overlap (5-78:75-145)

gi|116 MRGRVSPLMLLLGILVLASVSATHAKSSPYQKKTENPCAQRCLQSCQQEPDDLKQKACES
               10        20        30        40        50        60

                                10        20        30        40
Q41759           SAASPRG------RRAPVLHRALRRHPRHVRADDIRRHGRRDTVDARHLR
                    |||      :|:|  :|:  |:|     || ||: ||:   :|
gi|116 RCTKLEYDPRCVYDPRGHTGTTNQRSPPGERTRGRQPGDY--DDDRRQPRREE-GGRW--
               70        80        90       100         110

           50        60        70        80
Q41759 EHAPAPRREGRLRHLPRVSRQDTRRPPRDTQRPRFL
         :||  || : ::  |  |:| ||| :  |:||
gi|116 --GPAGPREREREEDWRQPREDWRRPSH--QQPRKIRPEGREGEQEWGTPGSHVREETSR
           120       130       140         150       160       170
```

Figure 4. Maize protein Q41759 peptide 503 Alignment against peanut allergen Ara h 1 obtained following analysis of 89 randomly selected corn proteins using a sliding window FASTA33 search. The alignment was observed exclusively with the sliding window search. The identity match occurred exactly at the 35% threshold and involved only 74 residues. The sliding window algorithm inserted six gaps and extended the effective alignment length, thus triggering the identity match.

potentially allergenic with the conventional FASTA analysis, while 53 54% were identified as allergenic using the sliding window analysis. The number of observed findings with the selected corn proteins unquestionably indicates a large number of false positives and a gross overestimation of the number of allergenic proteins in corn. These data are again likely attributed to the use of the stringent threshold of 35% and clearly represents an unrealistic view of the potential allergenic proteins in corn, which is a less common allergenic food. For example, if the threshold is raised to 50% as suggested by Aalberse [3] and Radauer and Breiteneder [4], the number of positive findings with the corn seed proteins is decreased by approximately half (17 and 25%, respectively) using either FASTA conventional or sliding window analysis. Increasing the threshold to 70% further diminishes the number of positive findings (3 and 6%, respectively) using either FASTA conventional or sliding window analysis, providing a more realistic estimate of allergenic proteins in corn [22, 23].

One finding with the randomly selected maize proteins that warrants additional discussion involves the identity match of the maize protein Q41759 (a hypothetical corn protein) to the peanut allergen Ara h 1. This alignment was observed exclusively with the sliding window search. The identity match occurred exactly at the 35% threshold and involved only 74 residues. Although less than 80 amino acids in length, the sliding window algorithm inserted 6 gaps and extended the effective alignment length, thus triggering the identity match (Fig. 4). The nature of this alignment and *E* score obtained (*i.e.*, 0.16), coupled with its absence in the conventional FASTA searches suggests that this identity match is an artifact/false positive resulting from the use of the sliding window algorithm.

The main empirical data to support the establishment of an identity threshold of 35% in 2001 came from a paper analyzing the apparent cross-reactivity of the birch pollen allergen Bet v 1 with proteins from cherry, apple, pear, celery, and celery [24] The relatively low level of amino acid sequence similarity observed, particularly when comparing the celery allergen Api g 1 and the carrot allergen Dau c 1 to Bet v 1 (~40% identity), in conjunction with reported cross-reactivity, served as the basis for the establishment of a threshold (*i.e.*, 35%) that would identify such relationships. Because of the role of the Bet v 1-like allergens played in defining the criteria currently recommended, a crucial benchmark for any *in silico* analysis would be to recognize and identify similarity between these allergens. Based on the analysis with Bet v 1 homologs, there was no difference in false positive rate observed between the conventional *versus* sliding window FASTA analyses. Therefore, the conventional FASTA algorithm is appropriate for detecting identities at or near the current recommended threshold of 35%. Interestingly, the lowest above threshold identity observed was 38% (Bet v 1 *vs.* Dau c 1; Table 5).

Recently, a nonallergenic α-amylase inhibitor from bean, when transformed into pea, displayed increased immunoreactivity in a nonvalidated animal model [6]. While this protein has generated much interest of late, it is not likely the protein would have been commercialized based on the current weight of evidence approach due to its observed identity to several allergenic lectins. Nevertheless, any modification of the FASTA analysis procedure should be evaluated against this protein to ensure that a positive match was returned. The bean α-amylase inhibitor protein was identified as a potential allergen using either the FASTA33 or 34 sliding window or conventional analysis. This *in silico* finding, however, would require further analysis and testing with sera from appropriate allergic patients to further investigate whether it would constitute a risk for individuals with specific allergies.

The inability to reliably identify any further target sequence from Ara h 1 in a putative nonallergenic test protein using the sliding window search *versus* a conventional search was for several reasons not unexpected. By using a sliding window, sequence is removed from the context of the entire protein. In the absence of the entire protein sequence, FASTA will insert gaps and generate a "globalized" alignment where sequence in the window is "stretched" to fit across the length of the database sequence. The overall impact of a sliding window is in some respects comparable to lowering the gap initiation and gap extension penalties. This was clearly demonstrated with the test sequence that contained the 15 amino acid spacing of the target sequences. In those instances that a window from the test protein aligned with Len c 1.0102, the alignment always included a 15 amino acid gap. Such globalized alignments are unlikely to reflect *bona fide* structural homology as they may be excessively gapped.

Searches using a sliding window will also tend to exaggerate the effect of low complexity regions on an alignment, as illustrated with the query sequences that generated multiple matches (Table 1). Low-complexity sequences yield alignments that are statistically significant but have little biological relevance. Although the Ara h 1 segment containing test proteins described herein did not contain any regions of low complexity, the use of criteria such as a 35% identity in 80 amino acids for the assessment of the significance of an alignment does not take advantage of the sophisticated statistical analyses (*i.e.*, a histogram of the identity scores and an *E* score) performed by the FASTA algorithm. These analyses include a histogram of the similarity scores and an *E* score. Inspection of the similarity histogram permits one to determine if the query sequence contains regions of low complexity. The *E* score is a statistical measure of the likelihood that the alignment is reliable. An *E* score of 1 or greater indicates that the alignment generated between the query and the database protein is no more meaningful than the alignment that would be obtained if the query sequence were shuffled prior to conducting the search. Therefore, the use of criteria such as alignment length (*i.e.*, 80 amino acid window), or combination of length and identity to select the most significant alignment in a series of sliding window searches is unreliable.

In summary, the data indicate that a conventional FASTA analysis compared to the sliding window analysis using the currently recommended threshold criteria of 35% or greater identity results in fewer potential false positive findings, while providing an equivalent false negative rate. The positive results obtained with the conventional FASTA analysis, however, still exceeded what would be predicted based on the expected percentage of real or true allergens in the clinic. This finding is likely attributed to the use of the currently recommended threshold criteria of 35%. For example, when the threshold was raised to 50% when evaluating corn seed protein sequences, the number of positive findings decreases by half using either the conventional or traditional FASTA analysis. In addition, the *E* values associated with the use of a conventional FASTA analysis were in general greater (<1) than those observed with the sliding window analysis and may suggest a more relevant identity to the query protein. Data further indicate that the use of criteria such as alignment length or a combination of length and identity to select the most significant alignment in a series of sliding window searches is unreliable. This is due to the following: (i) a sliding window search takes what should be a local alignment and makes it a global alignment by removing the sequence from the context of the entire protein. In the absence of the entire protein sequence, FASTA will insert gaps, and generate a "global" alignment where sequence in the window is stretched to fit across the length of the database sequence; (ii) the sliding window scoring regime does not take advantage of the statistical analysis performed by the FASTA algorithm (*i.e.*, a histogram of the similarity scores and an *E* score). Finally, the conventional FASTA analysis resulted in identity matches that better reflected functional similarities between proteins. In some cases, the sliding window analysis resulted in identity matches to a variety of proteins from different families with diverse functions. These data indicate that the 80 amino acid sliding window approach results in a greater number of potential false positive findings, as there appears to be little scientific justification for many of the matches (*i.e.*, matches occur between functionally divergent proteins). Therefore, it is recommended that the conventional FASTA analysis be conducted to compare the identity of a protein to known allergens.

5 References

[1] Food and Agriculture Organization, *Evaluation of Allergenicity of Genetically Modified Foods: Report of a Joint FAO/WHO Expert Consultation*, Rome 2001.

[2] Pearson, W. R., Lipman, D. J., Improved tools for biological comparison, *Proc. Natl. Acad. Sci. USA* 1988, *85*, 2440 2448.

[3] Aalberse, R. C., Structural biology of allergens, *J. Allergy Clin. Immunol.* 2000, *106*, 228 238.

[4] Radauer, C., Breiteneder, H., Pollen allergens are restricted to few protein families and show distinct patterns of species distribution, *J. Allergy Clin. Immunol.* 2006, *117*, 141 147.

[5] Baxevanis, A. D., Ouellette, B. F. F. (Eds.), *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*, John Wiley & Sons, Inc., New York 1998.

[6] Gendel, S. *Adv.* Sequence databases for assessing the potential allergenicity of proteins used in transgenic foods, *Food Nutr. Res.* 1998, *42*, 63 92.

[7] King, T. P., Hoffman, D., Lowenstein, H., Marsh, D. G., *et al.*, Allergen nomenclature, *Int. Arch. Allergy Immunol.* 1994, *105*, 224 233.

[8] Metcalfe, D. D., Astwood, J. D., Townsend, R., Sampson, H. A., *et al.* Assessment of the allergenic potential of foods derived from genetically engineered crop plants, *Crit. Rev. Food Sci. Nutr.* 1996, *36*, S165 S186.

[9] Prescott, V. E., Campbell, P. M., Moore, A., Mattes, J., *et al.*, Transgenic expression of bean α amylase inhibitor in peas results in altered structure and immunogenicity, *J. Agric. Food Chem.* 2005, *53*, 9023 9030.

[10] Reese, J. T., Pearson, W. R., Empirical determination of effec tive gap penalties for sequence comparison, *Bioinformatics* 2002, *18*, 1500 1507.

[11] Hefle, S. L., Nordlee, J. A., Taylor, S. L., Allergenic foods, *Crit. Rev. Food Sci. Nutri.* 1996, *36*, S69 S89.

[12] Moneret Vautrin, D. A., Kanny, G., Beaudouin, E., L'allergie alimentaire au mais existe t elle?, *Allerg. Immunol.* 1998, *30*, 230.

[13] Stadler, M. B., Stadler, B. M., Allergenicity prediction by protein sequence, *FASEB J.* 2003, *17*, 114 1143.

[14] Codex Alimentarius Commission, Alinorm 03/34: Joint FAO/ WHO Food Standard Programme, Codex Alimentarius Com mission, Twenty Fifth Session, Rome, Italy, June 30 July 5, Appendices III and IV, 2003, pp. 47 60.

[15] Breiteneder, H., Ebner, C., Molecular and biochemical classi fication of plant derived food allergens, C., *J. Allergy Clin. Immunol.* 2000, *106*, 27 36.

[16] Mills, E. N. C., Madsen, C., Shewry, P. R., Wichers, H. J., Food allergens of plant origin their molecular and evolution ary relationships, *Trends Food Sci. Technol.* 2003, *14*, 145 156.

[17] Mills, E. N. C., Jenkins, J. A., Alcocer, M. J. C., Shewry, P. R., Structural, biological, and evolutionary relationships of plant food allergens sensitizing via the gastrointestinal tract, *Crit. Rev. Food Sci. Nutri.* 2004, *44*, 379 407.

[18] Jenkins, J. A., Griffiths Jones, S., Shewry, P. R., Breiteneder, H., Mills, E. N. C., Structural relatedness of plant food aller gens with specific reference to cross reactive allergens: An analysis, *J. Allergy Clin. Immunol.* 2005, *115*, 163 170.

[19] Bredehorst, R., David, K. J., What establishes a protein as an allergen?, *Chromatogr. B. Biomed. Sci. Appl.* 2001, *756*, 33 40.

[20] Breiteneder, H., Mills, C., Structural bioinformatic approaches to understand cross reactivity, *Mol. Nutr. Food Res.* 2006, *50*, 628 632.

[21] Hileman, R. E., Silvanovich, A., Goodman, R. E., Rice, E. A. *et al.*, Bioinformatic Methods for Allergenicity Assessment Using a Comprehensive Allergen Database, *Int. Arch. Allergy Immunol.* 2002, *128*, 280 291.

[22] Pasini, G., Simonato, B., Curioni, A., Vincenzi, S., *et al.*, IgE mediated allergy to corn: a 50 kDa protein, belonging to the Reduced Soluble Proteins, is a major allergen, *Allergy* 2002, *57*, 98 106.

[23] Pastorello, E. A., Pompei, C., Pravettoni, V., Farioli, L., *et al.*, Lipid transfer protein is the major maize allergen maintain ing IgE binding activity after cooking at 100 degrees C, as demonstrated in anaphylactic patients and patients with posi tive double blind, placebo controlled food challenge results, *J. Allergy Clin. Immunol.* 2003, *112*, 775 783.

[24] Scheurer, S., Son, D. Y., Boehm, M., Karamloo, F., *et al.*, Cross reactivity and epitope analysis of Pru a 1, the major cherry allergen, *Mol. Immunol.* 1999, *36*, 155 167.

Molecular mechanisms in allergy and clinical immunology
(Supported by a grant from Merck & Co, Inc, West Point, Pa)

Series editor: Lanny J. Rosenwasser, MD

Structural biology of allergens

Rob C. Aalberse, PhD *Amsterdam, The Netherlands*

One of the major challenges of molecular allergy is to predict the allergenic potential of a protein, particularly in novel foods. Two aspects have to be distinguished: immunogenicity and cross-reactivity. Immunogenicity reflects the potential of a protein to induce IgE antibodies, whereas cross-reactivity is the reactivity of (usually preexisting) IgE antibodies with the target protein. In addition to these two issues, the relation between IgE-binding potential and clinical symptoms is of interest. This is influenced by physical properties (eg, stability and size) and immunologic properties (affinity and epitope valence). Discussions on immunogenicity and cross-reactivity of allergens rely on the establishment of structural similarities and differences among allergens and between allergens and nonallergens. For comparisons between the 3-dimensional protein folds, the representation as 2-dimensional proximity plots provides a convenient visual aid. Analysis of approximately 40 allergenic proteins (or parts of these proteins), of which the protein folds are either known or can be predicted on the basis of homology, indicates that most of these can be classified into 4 structural families: (1) antiparallel β-strands: the immunoglobulin-fold family (grass group 2, mite group 2), serine proteases (mite group 3, 6, and 9), and soybean-type trypsin inhibitor (Ole e 1, grass group 11); (2) antiparallel β-sheets intimately associated with one or more α-helices: tree group 1, lipocalin, profilin, aspartate protease (cockroach group 2); (3) (α+β) structures, in which the α- and β-structural elements are not intimately associated: mite group 1, lysozyme/lactalbumin, vespid group 5; and (4) α-helical: nonspecific lipid transfer protein, seed 2S protein, insect hemoglobin, fish parvalbumin, pollen calmodulin, mellitin from bee venom, Fel d 1 chain 1, serum albumin. Allergens with parallel β-strands (in combination with an α-helix linking the two strands, a motif commonly found in, for example, nucleotide-binding proteins) seem to be underrepresented. The conclusion is that allergens have no characteristic structural features other than that they need to be able to reach (and stimulate) immune cells and mast cells. Within this constraint, any antigen may be allergenic, particularly if it avoids activation of T_H2-suppressive mechanisms (CD8 cells and T_H1 cells). (J Allergy Clin Immunol 2000;106:228-38.)

Key words: *Allergen structure, allergenicity, cross-reactivity, epitope, protein folding, food allergen, novel foods, recombinant allergen*

Abbreviations used

CA:	α-Carbon
ER:	Endoplasmic reticulum
MnSOD:	Manganese-dependent superoxide dismutase
PDB:	Protein Data Bank
SCOP:	Structural Classification of Proteins

What makes an antigen an allergen? This question cannot be answered yet, but we are getting clues from different directions. Why is this a relevant question? First, there is the theoretic aspect: if we want to understand allergy, we need to understand allergens. There is a practical aspect as well: an answer to the question would help us decide whether the introduction of a new protein into our environment (particularly into our diet) increases the risk of subsequent allergic symptoms.[1-4] Moreover, modification or replacement of an allergenic protein may be of help for allergic patients, provided that the alternative protein (ie, the modified protein or the substitute) is less allergenic. Finally, if we understand allergens, we may find better ways of producing appropriately modified allergens for treatment.

Protein nomenclature has been confusing the understanding of protein structure for a long time. Proteins with the label *albumin* prove to have very different structures. Egg albumin, the primordial albumin, is structurally very different from serum albumin, milk albumin, fish parvalbumin, or seed 2S albumin. The term *pathogenesis-related protein*[5] is used for a series of plant proteins involved in stress reactions (comparable with the use of *acute-phase proteins* in mammals). Despite their common name, pathogenesis-related proteins have almost no structural relationship at all. Even in recent databases, a nomenclature issue has resulted in mislabeling a serum protein as an allergen. The rat serum prealbumin transthyretin was referred to as the major allergen Rat n 1 because it was labeled as prealbumin. However, Rat n 1 is a urinary prealbumin unrelated to serum prealbumin.

This review deals exclusively with protein allergens. IgE antibodies to nonpeptidic epitopes are known, for example, to classical haptens, such as the penicilloyl group,[6] and to glycosidic side chains of nonmammalian glycoproteins.[7] Protein structure can be described at different levels: primary structure (ie, the amino acid sequence), protein fold, domain structure, and surface structure. The surface structure is the most relevant for antibody binding, particularly the epitope, which is that part of the surface that on an atomic level interacts with the antibody.

From the Department of Allergy, CLB and the Laboratory for Experimental and Clinical Immunology, Academic Medical Center, University of Amsterdam, Amsterdam

Received for publication Apr 3, 2000; revised Apr 28, 2000; accepted for publication Apr 28, 2000

Reprint requests: Rob C Aalberse, PhD, the Department of Allergy, CLB and the Laboratory for Experimental and Clinical Immunology, Academic Medical Center, University of Amsterdam, PO Box 9190, 1006 AD, Amsterdam, The Netherlands

0091-6749/2000 $12 00 + 0 **1/1/108434**
doi:10 1067/mai 2000 108434

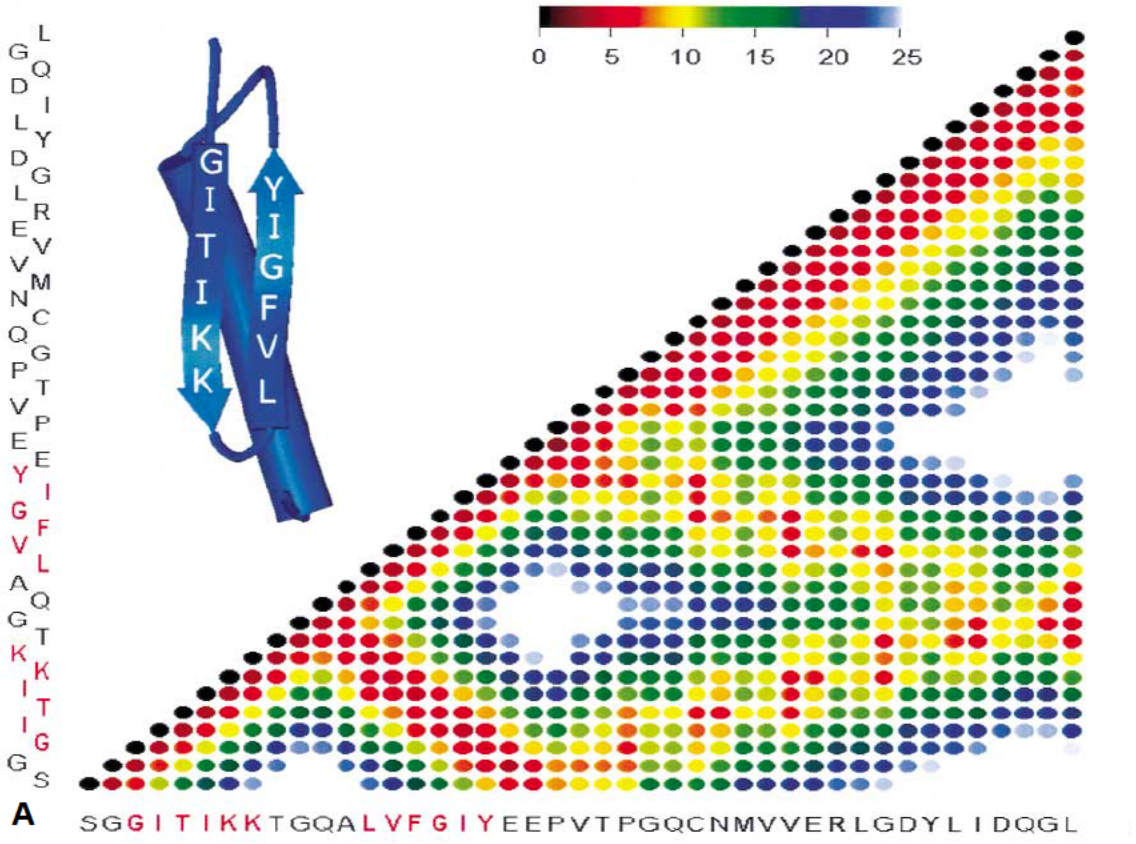


FIG 1. Two illustrations on the relationship between 3-dimensional structure and the corresponding proximity plots. The distances are color coded: red indicates a small distance, blue is a longer distance, and white is a distance of more than 2.5 nm. **A,** The C-terminal part of birch profilin (PDB code 1CQA), which consists of two antiparallel β-strands and an α-helix. The amino acids are indicated in their single-letter abbreviations from left to right and from bottom to top. *Red-colored dots* indicate that the two amino acids are at a short distance in the 3-dimensional structure; that is, the first amino acid of the first strand (GITIKK) is close to the last amino acid of the second strand (LVFGIY). As a start, it helps to focus on the distances of less than 1.0 nm (*red* or *orange dots*). The strands can be distinguished from the helix by counting the number of *red-orange dots* next to the diagonal: two for the strands and four for the helix. The antiparallel interactions between the two β-strands are visible as the series of adjacent *red dots* perpendicular to the diagonal. The interaction between the helix and each of the two β-strands produces a characteristic >-shaped checkered pattern; this checkered pattern results from the cyclic structure of the helix. **B,** The N-terminal part of schistosomal glutathione S-transferase (PDB code 1GTA). This fragment also contains two β-strands and one α-helix, but in this case the helix is between the two strands, and the strands are therefore parallel. The first amino acids in the first strand (ILGYW) are close to the first amino acids of the second strand (EEHLY). The interaction between the β-strands is visible in the lower-right corner as the series of adjacent *red squares* parallel to the diagonal (rather than perpendicular, as in **A**). This is a common fold, but it is relatively rare in allergens.

The current discussion will be focused on two aspects of allergen structure: a description (in general terms) of what is known about protein folds in allergens and about the characteristics of epitopes.

IMMUNOLOGIC VERSUS CLINICAL ALLERGENICITY

Semantically, the concept of allergenicity is ill defined. To a clinical allergist, allergenicity reflects the capacity of an antigen to induce symptoms or a skin reaction, whereas to an immunologist, it reflects either a peculiar type of immunogenicity (ie, the capacity of a protein to induce IgE antibodies) or simply the capacity to bind IgE antibodies.

Similarly, the term *allergen* is used to describe two or three distinct molecular properties: the property to sensitize (ie, induce the immune system to produce high-affinity antibodies, particularly of the IgE class) and the property to elicit an allergic reaction (ie, to trigger allergic symptoms in a sensitized subject). Moreover, it is also used to indicate the property to bind IgE antibodies. Complete allergens have all these properties. Some proteins, however, are known to elicit allergic symptoms but do not usually sensitize.

A well-known example of such a nonsensitizing elicitor (or incomplete allergen) is Mal d 1. This protein is the

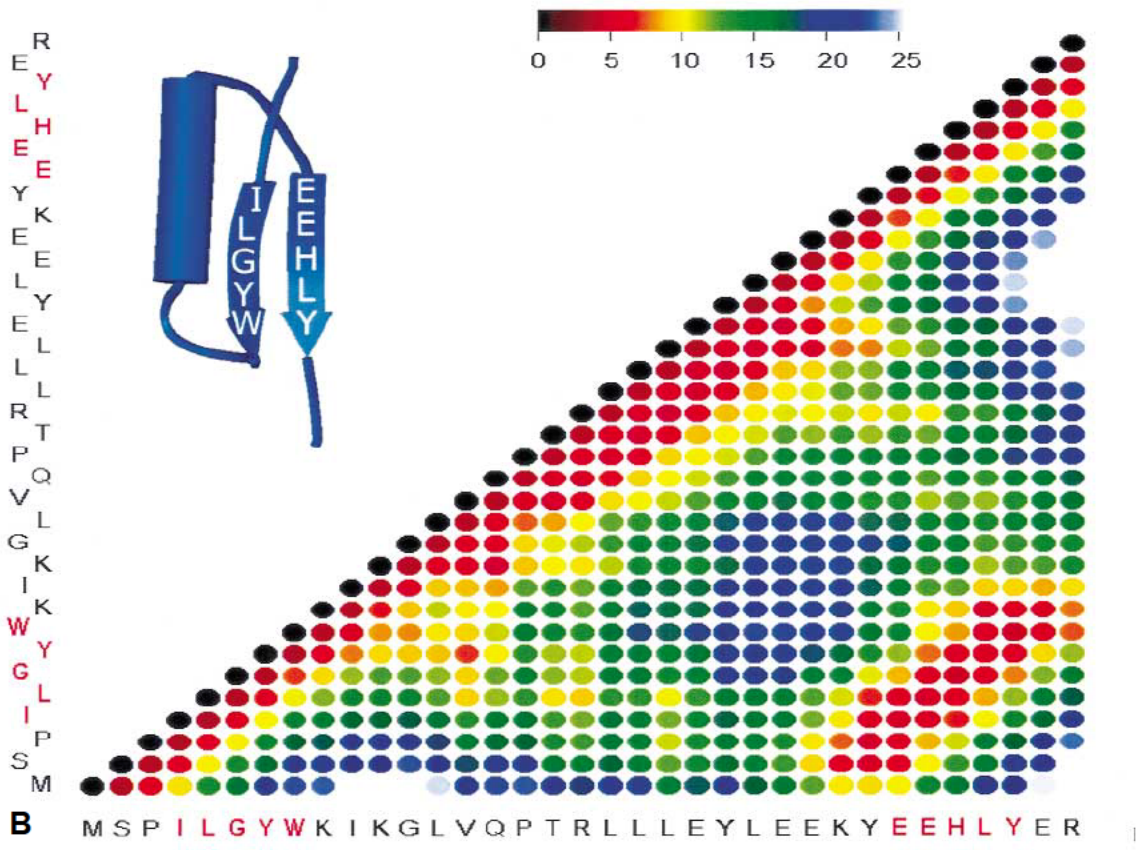


FIG 1. Continued.

apple homologue of the group 1 allergen from birch, Bet v 1. Inhalation of birch pollen induces IgE antibodies to Bet v 1, some of which cross-react with Mal d 1. Ingestion of apple does not induce IgE antibodies but may trigger activation of mast cells that are loaded with preexisting (birch-induced) cross-reactive IgE antibodies. By using RAST inhibition, the complete and incomplete allergens can be distinguished. In the example birch pollen will completely inhibit IgE binding to apple, whereas apple will give only partial inhibition of IgE binding to the birch allergen.

Obviously, oral exposure to Mal d 1 hardly, if ever, triggers the immune system to produce IgE antibodies, whereas inhalation of a very similar protein, Bet v 1, in quantities that are not vastly different often does. It seems likely that Mal d 1 does not have what it takes to make a protein a food allergen. One structural feature that may explain part of the differences between these incomplete food allergens and true food allergens, such as peanut allergens, is the susceptibility to the low pH in the stomach and to proteolytic enzymes, particularly pepsin.[8]

In this review allergenicity will be used in the immunogenicity sense: the induction of IgE antibodies. The most relevant second issue is allergic (in the clinical sense) reactivity of nonallergenic (ie, nonimmunogenic) proteins. This will be discussed under the heading of cross-reactivity.

WHEN IS A PROTEIN CONSIDERED TO BE A MAJOR ALLERGEN?

The distinction between major and minor allergens is relevant for various reasons but also in relation to the issue of allergenicity. The current definition of major allergen is based on the prevalence of IgE or skin reactivity in subjects that are sensitized (usually very strongly) to the total extract. This definition is unsatisfactory in that it does not reflect the contribution of the allergen to the overall reactivity of the extract. Intuitively, removal of a truly major allergen from an extract is expected to have a noticeable effect on the overall reactivity of that extract: a major allergen should make a difference. Such an interpretation invites a different type of definition. For example, a major allergen is responsible for more than 20% of the allergenic reactivity in more than 20% of the sensitized patients. This requires testing with extracts from which the allergen in question has been selectively removed (eg, with monospecific antibodies). Alternatively, it could be tested serologically by absorbing out all IgE antibodies to the allergen and then testing the residual activity of the absorbed serum. For most allergens, this aspect has not been studied. It is likely, however, that the major allergen claim made for many allergens would need to be reconsidered.

More relevant for the present discussion on allergenicity is to note that people outside the allergen-characteri-

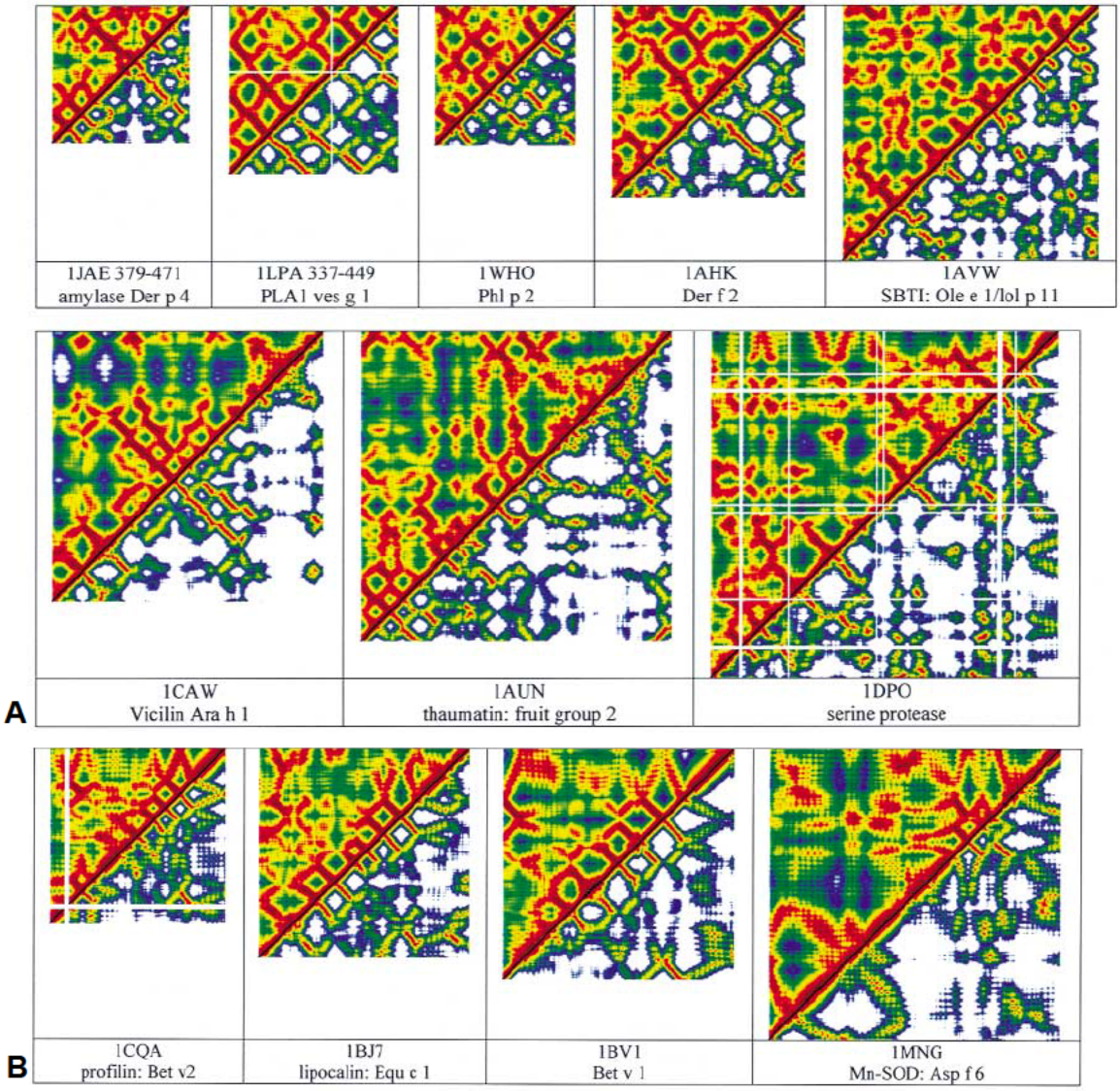


FIG 2. Classification of allergen folds. **A,** Group 1: anti-parallel β-strands. Note that the first two structures (1JAE 379-471 and 1LPA 337-449) are C-terminal domains of proteins. The complete proteins are shown in **E.** The antiparallel interaction between β-strands is indicated by the *red lines* perpendicular to the diagonal. Of vicilin (1CAW) and thaumatin (1AUN), only the N-terminal parts fall in this fold class. **B,** Group 2: antiparallel β-sheets intimately associated with one or more α-helices. The interaction between the β-sheet and the α-helix is indicated by the striped patterns caused by the periodicity of the helical structure. The C-terminal helix of Bet v 1 (1BV1) is further away from the β-sheet than the C-terminal helix of profilin (or lipocalin). Of superoxide dismutase (1MNG), only the C-terminal part falls in this class because the N-terminal part consists of two interacting helices. The two domains of trypsin (1DPO) can be seen to have a similar fold. **C,** Group 3: α+β structures, in which the α- and β-structural elements are not intimately associated. Note the striking similarity between lysozyme (1HEL) and lactalbumin (1HFZ).[40] **D,** Group 4: α-helical proteins. Interaction between helices can be recognized by the Scottish-kilt patterns caused by the periodicity of the helical structures. **E,** Larger proteins are plotted in a smaller format (25% of their original size). Note the repeating domain structures of serum albumin (3 similar domains) and ovotransferrin (2 similar domains). The *white lines* in some of the plots indicate gaps in the protein sequence.

zation field may assume that a major allergen is synonymous with a major allergenic risk. This is clearly not the case. For example, patatin has been reported to be a major allergen in potato[9] but is not a major allergenic risk, considering the frequency of exposure (ie, not as a potato protein but possibly as a latex allergen).[10,11] In terms of allergenicity, it would be relevant to compare the sensitizing potential of proteins for which the exposure is similar. In this context mite group 1 allergen would be a more major allergen (ie, a higher allergenic risk) than Can f 1, even though the latter is a major allergen. Some allergens are more major than others.

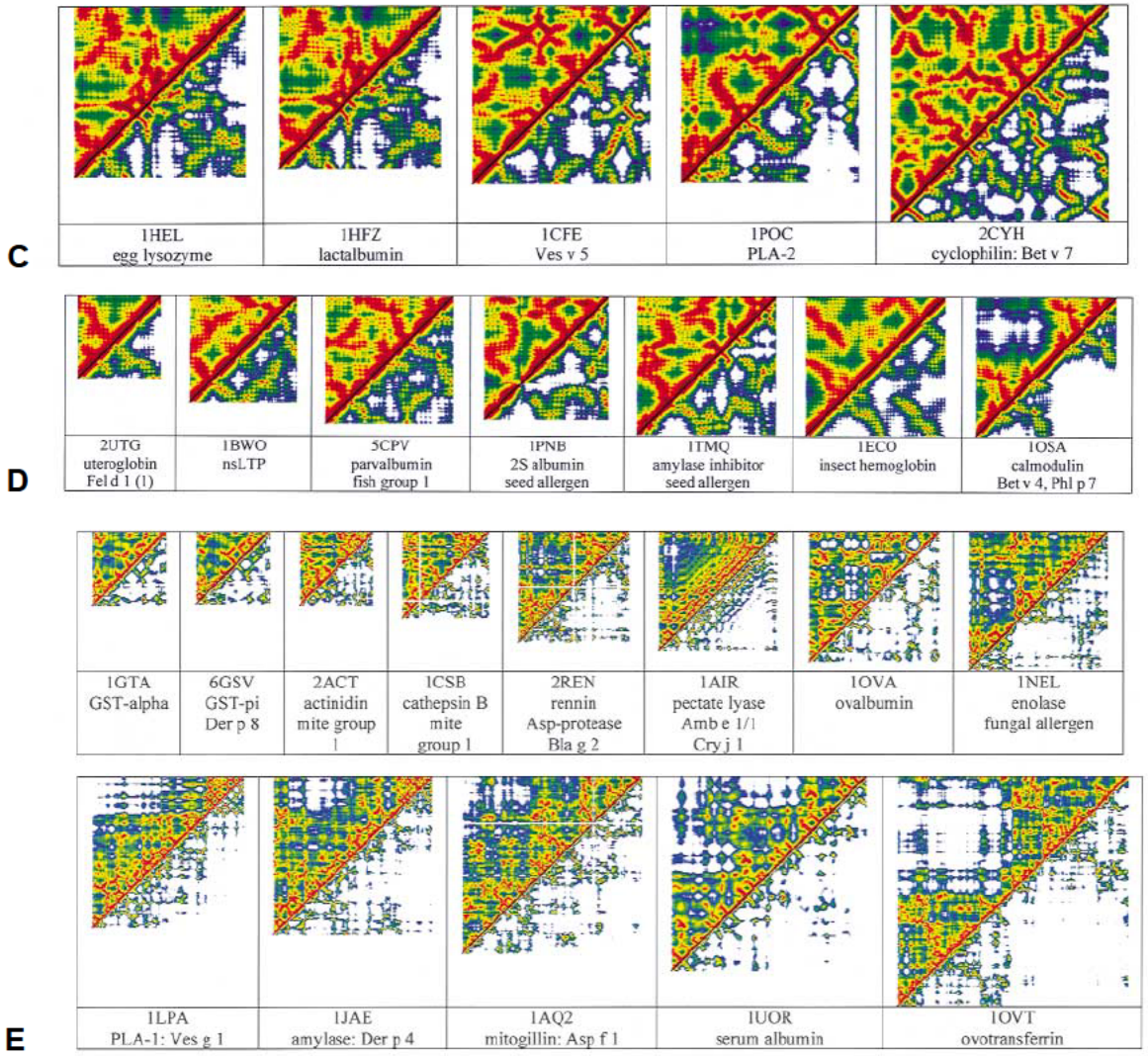


FIG 2. Continued.

DETERMINANTS OF ALLERGENICITY

Aspects of protein structure likely to be relevant for allergenicity are solubility, stability, size, and the compactness of the overall fold. These aspects reflect dependency of allergenicity on transport over mucosal barriers and susceptibility to proteases. Size and solubility of the intact protein would be relevant factors for airborne allergens more than for food allergens (for which limited proteolysis might enhance mucosal transport and hence allergenicity) or parenteral antigens, such as insect venoms, insect salivary allergens, invasive organisms (helminths and fungi), vaccines, or therapeutic proteins.

Posttranslational modification may affect allergenicity in different ways. It may induce new epitopes and it may affect solubility, stability, size, and susceptibility toward proteases. Moreover, uptake and processing by antigen-presenting cells are also known to be markedly influenced. Although glycosylation affects many of these processes, it is not a critical factor for allergenicity in general. Many allergens are not glycosylated, whereas some important allergens (eg, Gal d 1 [ovomucoid]) are heavily glycosylated.

It will become clear from this overview that few, if any, structural features are currently known to be common for allergens in general, even though most allergens can be grouped into a small number of structural classes. Some of these classes have been suggested to be intrinsically more allergenic, for example, 2S albumin from seeds[12] and lipocalin.[13] These will be discussed in more detail below.

It is, however, likely that features other than structure are more relevant for allergenicity. In addition, the search for common structural features relevant for allergenicity will become more relevant as we learn about the various pathways of allergic sensitization. It is not realistic to assume that the requirements for sensitization to food

TABLE I. Classification of protein folds in allergens

(1) Antiparallel β-strands
- The immunoglobulin-fold family
  - Grass group 2 (**1BMW, 1WHO, 1WHP**)
  - Grass group 1 (C-terminus)
  - Grass group 3
  - Mite group 2 (**1A9V**, **1AHK**, **1AHM**)
- Serine proteases (example: 1DPO, trypsin)
  - Mite group 3
  - Mite group 6
  - Mite group 9
- Soybean Kunitz-type trypsin inhibitor (1AVW)
  - Ole e 1
  - Grass group 11
- Fruits group 2: thaumatin (1AUN)
- Vicilin: peanut Ara h 1 (1CAW, 1DGR, 1DGW)

(2) Antiparallel β-sheets intimately associated with one or more α-helices
- Tree group 1 (**1BTV**, **1BV1**)
- Lipocalin
  - Milk β-lactoglobulin (**1BLG**)
  - Mouse (**1MUP**) and rat urinary protein (**2A2G**, **2A2U**)
  - Dog Can f 1
  - Dog Can f 2
  - Bovine Bos d 1
  - Horse Equ c 1 (**1BJ7**)
  - Cockroach Bla g 4
- Cystatin: cat allergen 4^{30} (1A67, 1CEW)
- Profilin (**1CQA**)
- Aspartate protease (2REN)
  - Cockroach Bla g 2

(3) (α+β) structures, in which the α- and β-structural elements are not intimately associated
- Mite group 1 (2ACT, 1CSB)
- Lysozyme (**1HEL**)/lactalbumin (**1HFZ**)
- Vespid group 5 (1CFE)
- Ovotransferrin = conalbumin (**1OVT**)
- Cyclophilin (2CYH)
  - Grass group 4
  - Tree group 7
- Phospholipase A2 (**1POC**)

(4) α-Helical
- Nonspecific lipid transfer protein (1BWO)
- Seed 2S albumin (1PNB)
- Insect hemoglobin (**1ECO**)
- Fish parvalbumin (**1CPD**, **5CPV**)
- Calmodulin (1OSA)
  - Bet v 4
  - Jun o 2
  - Phl p 7
- Mellitin from bee venom (**1MLT**)
- Fel d 1 chain 1 (2UTG)
- Serum albumin (1UOR)

(5) Other structures
- β-Helix: pectate lyase (1AIR, 2PEC)
  - Amb e 1
  - Amb e 2
  - Cry j 1
- Serine protease inhibitor (Serpin-family)
  - Ovalbumin (**1OVA**)
- PLA1 1LPA
- Glutathione S-transferase (1HNB, **1GTA**)
  - Cockroach group 5
- Mite group 8
  - Schistosomal glutathione S-transferase
- Mitogillin: Asp f 1 (**1AQ2**)
- MnSOD Asp f 6 (1MNG)
- Enolase (1NEL)
- Amylase (1JAE)
- Ovotransferrin (1OVT)
- Coiled coil: tropomyosin (1C1G, 1TMZ, 2TMA)
  - Shrimp group 1
  - Mite group 10
  - Cockroach
- Small proteins
- Ovomucoid (third domain only) 1OMU, 1OVO, 1CT4
- Hevein **1HEV**
- Amb e 5 **1BBG**, **2BBG**, **3BBG**

PDB codes printed in bold are coordinate files of the allergen itself; the other PDB files represent homology models

allergens in early childhood are very similar to those for late-onset sensitization to airborne occupational allergens.

DETERMINANTS OF CROSS-REACTIVITY

In contrast to allergenicity, cross-reactivity is largely determined by structural aspects: two proteins are cross-reactive only (almost) if they share structural features. What are the few exceptions to this rule?

Antibody affinity is an important consideration: low-affinity antibodies (particularly of the IgM class) have been found to be reactive with antigens with very little structural similarity. However, IgE-producing plasma cells evolved from a strongly T cell–dependent immune reaction in which only high-affinity B cells survive. Moreover, IgE-mediated cellular reactions are triggered by trace amounts of allergens. For these reasons, the discussion will be focused on high-affinity antibodies. It should be kept in mind, however, that information on the threshold of the affinity required for triggering allergic reactions by using cross-reactive allergens is still lacking. Another apparent exception is the cross-reactivity between glycoprotein allergens with unrelated protein structures.[7] In this situation cross-reactivity is due to shared glycans and thus similarity in structure.

Apart from these exceptions, all IgE cross-reactions described so far have been found to reflect shared features on the level of both primary and tertiary structure of the cross-reactive proteins. Whereas all cross-reactive proteins have a similar fold, the reverse is not true: proteins with a similar fold are not necessarily cross-reactive. This is partially due to immunologic tolerance induced by autologous proteins with a similar fold. More importantly, protein folding is liberal with respect to single amino acid substitutions for many positions in the sequence. Such substitutions may markedly affect the outer protein surface and thus reduce antibody reactivity. Similar protein folds are found with as little as 25%

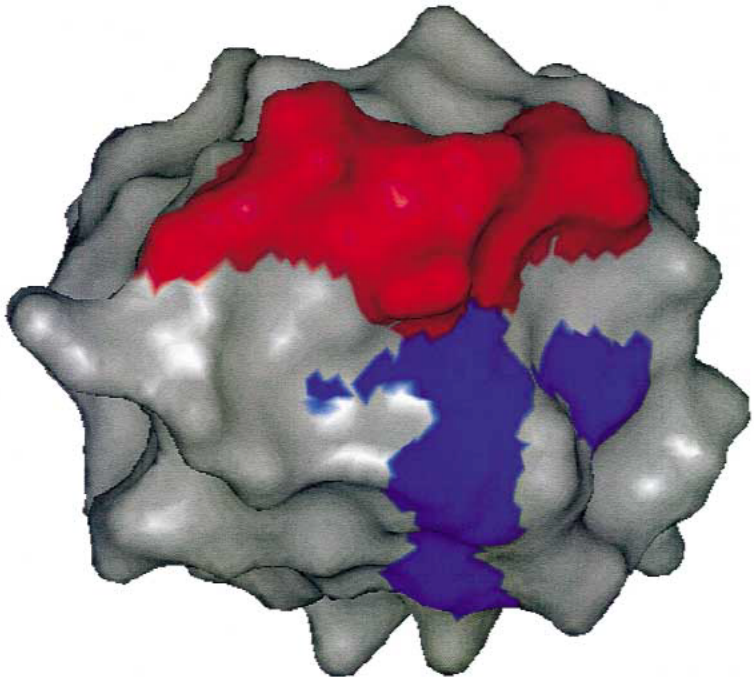


FIG 3. The epitope of egg white lysozyme as recognized by mAb D1.3 (PDB code 1FDL).[31] The amino acids that contribute to the N-terminal part of the epitope (18, 19, 22, 23, 24, and 27) are indicated in blue. The amino acids that contribute to the C-terminal part of the epitope (116-121, 124, and 125) are indicated in red. This figure was produced by using WebLabViewerPro (Molecular Simulations Inc).

amino acid identity, whereas cross-reactivity is rare below 50% identity. In most situations cross-reactivity requires more than 70% identity.

The likelihood of cross-reactivity depends not only on the similarity between the primary sequences of the two antigens. If a human homologue exists and is available to the immune system, this is likely to induce immune tolerance to the common epitopes. In some instances tolerance to the human protein appears to be absent, resulting in autoreactive IgE antibodies. The first description of such an autoreactive IgE antibody presumably induced by contact with a cross-reactive environmental allergen was the autoreactivity described for human profilin in patients with pollinosis.[14] Another striking example is the reactivity of IgE antibodies induced by the fungal allergen Asp f 6 (manganese-dependent superoxide dismutase [MnSOD]).[15,16] The recombinant human protein was found to give strongly positive skin reactions in subjects with reactivity to the *Aspergillus* protein.

To predict potential cross-reactivity of a new allergen (eg, an engineered food protein), it is informative to compare protein folds. In the absence of similarity in folding with allergens, protein cross-reactivity is virtually excluded (apart from the above-mentioned exceptions). If similarity in folding is observed, cross-reactivity needs to be investigated. The knowledge of the folding family allows a search restricted to subjects with sensitization to allergens with that particular protein fold. However, antibodies are highly idiotypical. Therefore absence of cross-reactivity of one antibody (or even many antibodies) does not exclude cross-reactivity of another antibody, even with the same epitope specificity. As with all risk assessments, statistical evaluation on the basis of a sufficiently large number of cases is important.

PROTEIN STRUCTURE

Many details on protein structure are given by Branden and Tooze.[17] The basis is the primary structure, the amino acid sequence. This information is becoming more easily obtainable through sequencing of complementary DNA. The amino acid sequence deduced from the nucleotide sequence is usually not the complete and final primary structure because cotranslational and posttranslational modifications are common. Most allergens are extracellular proteins. Secretion of a protein by a cell requires the presence of an N-terminal leader peptide for transport of the protein through the membrane of the endoplasmic reticulum (ER). This peptide is cleaved by a protease in the ER. In the ER other modifications occur, particularly glycosylation of asparagines, serines, and/or threonines. This glycosylation is actually a multistep process involving many glycosyl transferases and trimming glycosidases. These reactions often result in heterogeneity among the secreted glycoproteins. Particularly for plant glycoproteins, an important modification is the hydroxylation of prolines and lysines. These hydroxylated amino acids are also targets for glycosylation. The list of posttranslational modifications is

much longer and still growing, but this aspect of protein structure is beyond the scope of this review.

After in vitro unfolding, most proteins will fold at least partially back into their native structure. Therefore the primary sequence contains all the clues for the 3-dimensional structure, and we should be able to predict the 3-dimensional structure of the protein. In practice this is still a distant dream. However, progress in this field is rapid because of the rapidly increasing number of 3-dimensional structures that have been and are being solved and the increased computational power. Homology modeling proves to be reliable for sequences that can be aligned well. With sequence homology as low as 25%, overall backbone folds can be predicted often with high accuracy, with the exception of some of the more complicated loops. Configurations of solvent-exposed side chains are far more difficult to predict, and therefore even if the 3-dimensional structure of two proteins with similar folds is known, prediction of cross-reactivity is not yet possible. Fold recognition is, however, an important help in the search for potential cross-reactivity.

Fold recognition on the basis of the primary sequence is dependent on the classification of folds, which can be automated to a certain extent,[18-24] but remains to some degree a subjective process. For this review, I have used the Structural Classification of Proteins (SCOP).[25] To visually compare the folds of allergens, these folds are represented in the format of a proximity matrix.[26] The rationale for this procedure is described below.

PROXIMITY PLOT: COMPARING PROTEIN FOLDS AND DOMAIN INTERACTIONS

The structural basis of a folded protein is its peptide backbone. This consists of linear repeats of the 3 atoms (the amino-nitrogen, the α-carbon, and the carbonyl carbon) that form the basic peptide triplet $(\text{-N-CA-CO-})_n$. Because the goniometric angles between the atoms are fixed, the distance between two neighboring α-carbons (CAs) is virtually constant (0.38 nm). The degree of contraction of the protein backbone largely depends on torsion angles (ie, the angles along the axis between two atoms). The two most relevant torsion angles are torsion along the N-CA axis (Phi) and torsion along the CA-CO-axis (Psi). In its most extended form (the β-strand), the length of the backbone increases by approximately 0.33 nm for each subsequent amino acid. In its contracted form (ie, the well-known α-helix), the length increases by approximately half this distance. Intermediates between these two extremes (the extended β-strand and the contracted α-helix) are referred to as turns or coils (depending, among others, on the number of amino acids involved). These secondary structure elements are stabilized by hydrogen bonds between β-strands (resulting in β-sheets) or within α-helices, turns, or coils.

Interactions between these elements of secondary structure results in the tertiary structure: domains of usually 80 to 200 amino acids often stabilized by intrachain disulfide bonds. In many enzymes two domains interact to form a substrate-binding cleft.

These 3-dimensional structures are usually presented in the form of simplified cartoons. This works particularly well on a computer screen with programs such as RASMOL (see Appendix). However, a visual comparison between more than two such 3-dimensional structures is unrewarding, partially because of the limitations of the visual system, partially also because of the dependency on the rotation of the protein structures that are to be compared.

The information present in the x, y, and z coordinates (which are the basis of the structure in space) can be presented as a rotation-independent 2-dimensional plot by using distances between the CAs of the amino acids rather than their actual position. Because secondary and tertiary structures depend on contacts between amino acids, the usual 2-dimensional representation of these distances emphasizes proximity rather than distance: the shorter the distance between 2 amino acids, the more emphasis. Hence the use of the term *proximity plot*.

For a protein of 100 amino acids, the 10,000 distances between the CAs are calculated and presented as a symmetrical square (100×100) matrix. For the diagonal, the distances are obviously zero. All the values next to the diagonal are also unrevealing because these are all 0.38 nm (see above). The values next off the diagonal are already more interesting because values of more than 0.65 nm indicate a β-strand, whereas values smaller than 0.55 nm indicate an α-helix (or a turn). Cells in the matrix that are close to the diagonal thus provide information on the secondary structure. Low values (ie, small distances) in cells that are away from the diagonal indicate contact between elements of secondary structure; that is, they provide information on the tertiary structure (or in multichain structures on interactions between protein modules). Such a distance matrix contains all the information necessary to rebuild the 3-dimensional structure (apart from chirality because the mirror image of the protein will give the same distance matrix). For easy viewing, the distances relative to some cut-off values are color coded and plotted. In the examples presented here, cut-off values of 5.0 and 2.5 nm have been used.

One of the convenient aspects of these proximity plots is that they enable a quick, objective, and rotation-independent comparison between protein folds. In Fig 1 two examples are shown in more detail to illustrate the basic principles. Both consist of protein fragments containing two β-strands and one α-helix connected in a simple up-and-down fashion. In the first example (the C terminus of birch profilin[27]) the sequence is β-β-α, which implies that the β-strands run antiparallel. In the second example (the N-terminus of glutathione S-transferase[28]) the sequence is β-α-β, and therefore the β-strands run in parallel.

Even if the structural interpretation of these proximity plots (Fig 2) requires some practice, one message can easily be derived from these plots by using them as fingerprints: the more different the plots, the more different the folds. Comparisons on the basis of the plots are liberal, and therefore proteins with similar plots may have significant differences in folding. However, if two proteins have different plots, they definitely have different folds.

CLASSIFICATION OF PROTEIN FOLDS IN ALLERGENS

Technical details can be found in the Appendix. The table contains a selection of allergens (fully listed on the World Wide Web at www.allergen.org/List.htm) for which the structure has been determined directly or for which the structure can be predicted with a highly variable degree of accuracy from the primary sequence by homology with proteins with a known structure. The literature references for these structures can be found in the Protein Data Bank (PDB) files.[29]

Whereas the basis for the classification in Table I[30] was provided by the SCOP database,[19] which is also an excellent source for finding PDB codes, the classification used was slightly adapted to combine allergens with similar features into a single class. The main issue was the group of proteins with one or two prominent antiparallel β-sheets in conjunction with one or two α-helices. Depending on the size of the helix, SCOP classifies some allergens as β (lipocalin and aspartate protease) and others as α+β (Bet v 1, profilin, and cystatin).

Some structures occur infrequently in the allergens investigated. Some strikingly unique allergens are tropomyosin (a long, filamentous, coiled, coil protein very different from the usual globular structures found for most allergens) and the allergens belonging to the pectate lyase family (eg, Amb e 1 and Cry j 1). The structure for this group of proteins is a long series of β-strands wound into a helix in which these β-strands form parallel sheets.

Many allergens are small, single-domain structures; some are so small that classification is problematic (Amb e 5 and hevein). However, complex multifold allergens do exist, particularly among the allergens that do not have to pass the airway mucosal barrier. Examples are the fungal proteins mitogillin, enolase, and MnSOD and venom phospholipase A.

EPITOPES: LINEAR OR CONFORMATIONAL?

The part of the surface of the allergen that interacts with the Fab fragment of the antibody (ie, the epitope) is of particular interest.

The atomic details of the interaction between antibody and antigen are known for more than 30 antibody-antigen combinations. The interaction between two proteins is usually described in terms of buried surface (ie, the surface area that becomes inaccessible to water molecules as a consequence of the formation of the immune complex). This value is calculated from crystallographic data by using a spherical probe the size of a water molecule. The review by Padlan[31] gives a range of 54.0 to 89.0 nm^2 for the immune complexes with intact protein antigens (as opposed to peptides or haptens) that were available at that time. For a typical globular 20-kd allergen, the radius is approximately 2.0 nm, with a spherical surface area of 500.0 nm^2. The water-accessible surface area varies depending on the compactness of the folded structure but will usually be approximately twice this value or 1000.0 nm^2 in this example. Thus the interaction with a Fab fragment of an antibody will bury approximately 5% to 10% of the surface of such a 20-kd allergen. The surface area that is accessible to a water molecule is obviously larger than the surface area accessible to a Fab fragment (and even more so for a complete antibody). It is therefore likely that a single allergen of this size cannot accommodate more than 5 to 10 antibodies at the same time.

Before crystallographic information was available, information on the structure of epitopes was derived, among others, from the reactivity of the antibody with peptides derived from the linear sequence of the allergen or with the unfolded allergen on an immunoblot. On the basis of this type of information, epitopes were classified as either linear or conformational. An antibody was assumed to react with a linear epitope if it was reactive with a small peptide (usually 8-15 amino acids) derived from the linear sequence or if it was reactive (in an immunoblot) after the allergen had been unfolded by boiling in SDS in the presence of a reducing agent. However, from the crystallographic studies, a general picture emerges that does not support the dichotomy of linear versus conformational epitopes: all protein epitopes are conformational. In all examples studied this buried surface area involves amino acids from different parts of the linear sequence; that is, all these epitopes are conformational to some extent. Fig 3 shows the epitope of lysozyme for mAb D1.3.[32]

How can this observation be reconciled with the observation that many antibodies react with small peptides derived from the linear sequence of the antigen? The main factor is the huge difference in affinity. On a molar basis, the peptides are very inefficient compared with the intact antigen, inhibiting the interaction between antigen and antibody. The peptide represents a fraction of the epitope, and the strength of the interaction with the antibody is a fraction of the strength with the complete antigen. The strength of interaction with the peptide is usually decreased even more because of the higher flexibility of the free peptide compared with the peptide in the complete antigen, resulting in a larger loss of entropy on binding. For the interpretation of the immunoblot data, it is relevant to appreciate that many blotted proteins refold extensively on removal of the SDS, particularly if reoxidation of any disulfides that might be present is allowed.

CONCLUSIONS

The information on the atomic details of allergen structures indicates that allergens are heterogeneous, also from a structural point of view. Even if some folds are less prevalent among the currently known allergens, none of the protein folds seem to be incompatible with allergenicity. The current review is, however, biased toward allergens with a well-defined, stable structure.

What then determines allergenicity? Why are some allergens more major than others? The level and route of

exposure are obviously relevant. Similarity to human homologous proteins needs to be considered, but this relation is likely to be complex (depending on the accessibility of the human protein to the tolerizing mechanisms of the immune system). For food allergens, digestibility is clearly a factor to consider. For airborne allergens, size and solubility are important.

The high allergenicity of castor bean dust is presumably due to a negative adjuvant function of one of its components, the antisuppressive activities of ricin.[33] The IgE-inducing potential of helminths is likely to reflect similar mechanisms. Inhaled proteases, particularly mite group 1, may act in the same way[34-36] or might enhance the local permeability of the airway mucosa.[37-39] Because most allergens do not possess protease activity, it would be important to investigate whether inhaled mite group 1 allergens have an IgE-promoting adjuvant effect. Because the prevalence of allergy is not strikingly lower in areas with low mite numbers, the case for an obligatory connection between enzymatic activity and allergenicity is as yet unconvincing.

We clearly need more information. However, a shift in focus may be needed. The main lack of information is not on major allergens but on their counterparts, the nonallergenic antigens. We need to identify and characterize antigens that are able to reach (and stimulate) immune cells and yet are not allergenic. The current data suggest the following working hypothesis: within the constraint mentioned, allergenicity depends on a single specific property to avoid activation of T_H2-suppressive mechanisms.

I thank Ronald van Ree and Kaare Meno for helpful discussions. Ed Nieuwenhuys, Walter Langhorst, and Piet Schrama have been helping out whenever the computers did not do what I had in mind.

REFERENCES

1 Lehrer SB, Horner WE, Reese G Why are some proteins allergenic? Implications for biotechnology Crit Rev Food Sci Nutr 1996;36:553-64
2 Kimber I, Kerkvliet NI, Taylor SL, Astwood JD, Sarlo K, Dearman RJ Toxicology of protein allergenicity: prediction and characterization Toxicol Sci 1999;48:157-62
3 Metcalfe DD, Astwood JD, Townsend R, Sampson HA, Taylor SL, Fuchs RL Assessment of the allergenic potential of foods derived from genetically engineered crop plants Crit Rev Food Sci Nutr 1996;36(Suppl):S165-86
4 Astwood JD, Fuchs RL Allergenicity of foods derived from transgenic plants Monogr Allergy 1996;32:105-20
5 Van Loon LC, Van Strien EA The families of pathogenesis-related protein, their activities, and comparative analysis of PR-1 types proteins Physiol Mol Plant Pathol 1999;55:85-97
6 Baldo BA Penicillins and cephalosporins as allergens—structural aspects of recognition and cross-reactions Clin Exp Allergy 1999;29:744-9
7 Aalberse RC, Van Ree R Crossreactive carbohydrate determinants Clin Rev Allergy Immunol 1997;15:375-87
8 Astwood JD, Leach JN, Fuchs RL Stability of food allergens to digestion in vitro Nat Biotechnol 1996;14:1269-73
9 Seppala U, Alenius H, Turjanmaa K, Reunala T, Palosuo T, Kalkkinen N Identification of patatin as a novel allergen for children with positive skin prick test responses to raw potato J Allergy Clin Immunol 1999;103:165-71
10 Kostyal DA, Hickey VL, Noti JD, Sussman GL, Beezhold DH Cloning and characterization of a latex allergen (Hev b 7): homology to patatin, a plant PLA2 Clin Exp Immunol 1998;112:355-62
11 Breiteneder H, Sowka S, Wagner S, Krebitz M, Hafner C, Kinaciyan T, et al Cloning of the patatin-like latex allergen Hev b 7, its expression in the yeast *Pichia pastoris* and its immunological characterization Int Arch Allergy Immunol 1999;118:309-10
12 Teuber SS, Dandekar AM, Peterson WR, Sellers CL Cloning and sequencing of a gene encoding a 2S albumin seed storage protein precursor from English walnut *(Juglans regia)*, a major food allergen J Allergy Clin Immunol 1998;101:807-14
13 Virtanen T, Zeiler T, Mantyjarvi R Important animal allergens are lipocalin proteins: why are they allergenic? Int Arch Allergy Immunol 1999;120:247-58
14 Valenta R, Duchene M, Pettenburger K, Sillaber C, Valent P, Bettelheim P, et al Identification of profilin as a novel pollen allergen; IgE autoreactivity in sensitized individuals Science 1991;253:557-60
15 Mayer C, Appenzeller U, Seelbach H, Achatz G, Oberkofler H, Breitenbach M, et al Humoral and cell-mediated autoimmune reactions to human acidic ribosomal P2 protein in individuals sensitized to *Aspergillus fumigatus* P2 protein J Exp Med 1999;189:1507-12
16 Crameri R, Faith A, Hemmann S, Jaussi R, Ismail C, Menz G, et al Humoral and cell-mediated autoimmunity in allergy to *Aspergillus fumigatus* J Exp Med 1996;184:265-70
17 Branden C, Tooze J Introduxtion to protein structure 2nd ed New York: Garland Publishing; 1999
18 Holm L, Sander C Protein folds and families: sequence and structure alignments Nucleic Acids Res 1999;27:244-7
19 Hubbard TJ, Ailey B, Brenner SE, Murzin AG, Chothia C SCOP: a structural classification of proteins database Nucleic Acids Res 1999;27:254-6
20 Aalberse RC, Van der Zee JS, Vlug A IgG4 antibodies in atopic allergy Lab Management 1985;23:19-32
21 Pietrokovski S, Henikoff JG, Henikoff S The Blocks database—a system for protein classification Nucleic Acids Res 1996;24:197-200
22 Corpet F, Gouzy J, Kahn D Recent improvements of the ProDom database of protein domain families Nucleic Acids Res 1999;27:263-7
23 Holm L, Sander C Protein folds and families: sequence and structure alignments Nucleic Acids Res 1999;27:244-7
24 Holm L, Sander C Dictionary of recurrent domains in protein structures Proteins 1998;33:88-96
25 Seiberler S, Natter S, Hufnagl P, Binder BR, Valenta R Characterization of IgE-reactive autoantigens in atopic dermatitis 2 A pilot study on IgE versus IgG subclass response and seasonal variation of IgE autoreactivity Int Arch Allergy Immunol 1999;120:117-25
26 Holm L, Sander C Protein structure comparison by alignment of distance matrices J Mol Biol 1993;233:123-38
27 Fedorov AA, Ball T, Mahoney NM, Valenta R, Almo SC The molecular basis for allergen cross-reactivity: crystal structure and IgE-epitope mapping of birch pollen profilin Structure 1997;5:33-45
28 McTigue MA, Bernstein SL, Williams DR, Tainer JA Purification and crystallization of a schistosomal glutathione S-transferase Proteins 1995;22:55-7
29 Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al The Protein Data Bank Nucleic Acids Res 2000;28:235-42
30 Chapman MD, Ichikawa K, Vailes LD Molecular cloning of a new cat allergen, cystatin, a cysteine protease inhibitor [abstract] J Allergy Clin Immunol 1999;103:S122
31 Padlan EA X-ray crystallography of antibodies Adv Protein Chem 1996;49:57-133
32 Fischmann TO, Bentley GA, Bhat TN, Boulot G, Mariuzza RA, Phillips SE, et al Crystallographic refinement of the three-dimensional structure of the FabD1 3-lysozyme complex at 2 5-A resolution J Biol Chem 1991;266:12915-20
33 Underwood SL, Kemeny DM, Lee TH, Raeburn D, Karlsson JA IgE production, antigen-induced airway inflammation and airway hyperreactivity in the brown Norway rat: the effects of ricin Immunology 1995; 85:256-61
34 Shakib F, Schulz O, Sewell H A mite subversive: cleavage of CD23 and CD25 by Der p 1 enhances allergenicity Immunol Today 1998;19:313-6
35 Schulz O, Sewell HF, Shakib F The interaction between the dust mite antigen Der p 1 and cell-signalling molecules in amplifying allergic disease Clin Exp Allergy 1999;29:439-44
36 Gough L, Schulz O, Sewell HF, Shakib F The cysteine protease activity of the major dust mite allergen Der p 1 selectively enhances the immunoglobulin E antibody response J Exp Med 1999;190:1897-902
37 Robinson C, Kalsheker NA, Srinivasan N, King CM, Garrod DR, Thompson PJ, et al On the potential significance of the enzymatic activ-

ity of mite allergens to immunogenicity Clues to structure and function revealed by molecular characterization Clin Exp Allergy 1997;27:10-21
38 Stewart GA Dust mite allergens Clin Rev Allergy Immunol 1995;13:135-50
39 Wan H, Winton HL, Soeller C, Tovey ER, Gruenert DC, Thompson PJ, et al Der p 1 facilitates transepithelial allergen delivery by disruption of tight junctions [see comments] J Clin Invest 1999;104:123-33
40 Smith SG, Lewis M, Aschaffenburg R, Fenna RE, Wilson IA, Sundaralingam M, et al Crystallographic analysis of the three-dimensional structure of baboon alpha-lactalbumin at low resolution Homology with lysozyme Biochem J 1987;242:353-60

APPENDIX
Technical considerations

The 4-character PDB code of the protein structure gives access to the file that contains the coordinates of most of the atoms in the structure. For example, one of the files for the birch allergen Bet v 1 has the PDB code 1BV1. The complete filename is "PDB1BV1.ENT". The PDB[29] can be freely accessed through the Internet via SearchLite at http://www.rcsb.org/pdb/searchlite.html. Coordinate files, such as PDB1BV1.ENT, can be downloaded and opened in either a text viewer by using a text editor, such as WORDPAD, or a word processor. They can also be opened in a molecular viewer, such as RASMOL, which is also freely available through the Internet at http://www.umass.edu/microbio/rasmol/. The PDB file contains not only the x, y, and z coordinates for each atom in Ångstroms (1 Ångstrom = 0.1 nm), but also background information, including the authors and literature references.

SCOP is accessible through http://scop.mrc-lmb.cam.ac.uk/scop/.

The PDB files contain the information required to produce the proximity plots for which an option is available in several molecular modeling packages. The plots in this review have been produced by using PROXIMA.EXE, a small (51 kB) DOS-based program, available for noncommercial use from the author by means of e-mail (aalberse@clb.nl).

The FASEB Journal express article 10.1096/fj.02-1052fje. Published online April 22, 2003.

Allergenicity prediction by protein sequence

Michael B. Stadler and Beda M. Stadler

Institute of Immunology, University of Bern, Switzerland

Corresponding author: Michael B. Stadler, Institute of Immunology, Sahlihaus 2, InselspitalCH-3010 Bern, Switzerland. E-mail:michael.stadler@insel.ch

ABSTRACT

Potential allergenicity of transgenic proteins for consumption must be investigated before their introduction into the food chain. A prerequisite is sequence analysis. We have critically reviewed the performance of the current guidelines proposed by the Food and Agriculture Organization (FAO) and the World Health Organization (WHO) for allergenicity prediction based on protein sequence and show that its precision is very low. To improve prediction, we propose a new strategy based on sequence motifs identified from a new allergen database. If tested on random test sequences and known allergens, both methods are apparently very sensitive. However, the precision of our motif-based prediction (95.5%) is superior to the current method (36.6%). We conclude that the proposed motif-based prediction is a superior alternative to the current method for use in the decision-tree approach for allergenicity assessment.

Key words: genetically modified food • allergen database • allergenicity prediction • bioinformatics • sequence motifs

Allergens inducing type I allergic responses are proteins that elicit specific IgE antibodies. The allergic reaction is triggered by allergens aggregating IgE antibodies bound to high-affinity Fc receptor (FcεRI) on mast cells and basophils (1, 2). Mediators released by activated cells cause the symptoms of allergy, such as sneezing and swelling of the mucosa, characteristic for allergic rhinitis, allergic conjunctivitis, and asthma.

Despite the great number of presently identified allergenic proteins (3), it is still not known why only few and particular proteins that humans are exposed to provoke allergic reactions. Thus, a method for allergenicity prediction would be beneficial, especially in order to prevent the inadvertent generation of new allergenic food plants by agricultural biotechnology.

In 1996, a task force of the International Food Biotechnology Council (IFBC) and the Allergy and Immunology Institute of the International Life Sciences Institute (ILSI) developed a decision-tree approach for the assessment of potential allergenicity of plants produced through agricultural biotechnology (4, 5). In 2001, FAO and WHO have modified the approach in a joint expert consultation on foods derived from biotechnology. In the consultation report (accessible at http://www.who.int/fsf/GMfood/), guidelines have been published for the evaluation of allergenicity of genetically modified foods. Besides biological tests concerning the protein of interest, a standard method for sequence comparison has been defined. Briefly, a protein is

considered allergenic if it shares more than 35% sequence similarity (window of 80 residues) or an identity of at least six contiguous amino acids with a known allergen.

In this study, we critically review the performance of the proposed method for the prediction based on sequence similarity. We present an automated approach for the construction and update of a local allergen sequence database. Using probabilistic sequence motifs identified from this allergen database, we propose a new approach for allergenicity prediction in order to overcome the low precision of the current method.

MATERIALS AND METHODS

Databases and Software

External sequence databases downloaded and installed locally to be used for the study: Swiss-Prot (6): Release 40.0; 101,602 proteins; obtained from: ftp://ftp.expasy.org/databases/swiss-prot. Randomized Swiss-Prot: Sequences of release 40 shuffled in consecutive windows of 20 amino acids. trGEN (7) Human, Release 12–19–2001; 330,743 sequences, obtained from ftp://ftp.isrec.isb-sib.ch/pub/databases/trgen/. Swiss-Prot allergen index: Release 16-Oct-2001; 274 protein sequences; available at: http://www.expasy.org/cgi-bin/lists?allergen.txt. Rice (8, 9): TIGR rice gene index (OsGI); rRelease 7.0; 10,891 protein sequences, obtained from: http://www.tigr.org/tdb/ogi. Remote allergen lists used for allergen database construction: http://www.allergen.org (10), http://www.expasy.org/cgi-bin/lists?allergen.txt (6), http://www.iit.edu/~sgendel (11). Remote sequence databases used as a source for allergen sequences: GenBank (12). PIR (13): Online access at: http://www.ncbi.nlm.nih.gov/entrez. A local allergen database was generated by extracting all accession numbers in the published allergen lists and downloading the corresponding sequences from the public sequence databases. Subsequently, DNA sequences were translated, and sequence variants were generated according to annotation. Finally, all redundant sequences were removed, resulting in a database containing 779 allergens (February 11, 2002). Freely available software packages downloaded and installed locally to be used for the study: pftools (14), Version 2.2, obtained from http://www.isrec.isb-sib.ch/ftp-server/pftools. MEME (15): Version 3.0.3, obtained from http://meme.sdsc.edu/meme/website. FASTA (16): Version 3.4, obtained from ftp://ftp.virginia.edu/pub/fasta. NCBI-BLAST (17): Version 2.2.1, obtained from ftp://ftp.ncbi.nih.gov/blast/. Scripts generating the allergen database and controlling iterative motif discovery and allergenicity prediction were written in Perl (http://www.perl.org) using extensions from Bioperl (http://www.bioperl.org) for sequence processing and online data retrieval.

FAO/WHO allergenicity evaluation

According to the guidelines for allergenicity evaluation of foods derived from biotechnology (full report at http://www.who.int/fsf/GMfood/), a query protein is potentially allergenic if it either has an identity of at least six contiguous amino acids or more than 35% sequence similarity over a window of 80 amino acids when compared with a known allergen. We have written a program that compares a query protein with each allergen and rates it allergenic, if

either of the two criteria were fulfilled. The value for identity length *n* could be specified as a parameter to allow for more flexible testing.

For allergen prediction in the allergen database, we slightly modified the program by removing the single-query allergen sequence from the reference allergen sequence database. Without this modification, each query sequence would be contained in the allergen database, and identical subsequences of *n* residues could always be found (for *n* not greater than sequence length).

Automated iterative allergen motif discovery

Starting with all 779 sequences in the allergen database, the following steps were performed iteratively until no motif with E-value less than 0.01 could be identified: MEME (15) (*zoops* motif match mode) was used to identify the most relevant motif of 50 residues contained in the allergen sequences. The length of 50 residues was chosen to be shorter than the mean length of a protein domain in order to prevent generation of multi-domain motifs. The mean domain length of the 974'587 Pfam (18) domains identified in Swiss-Prot and TrEMBL is 135 residues, and 79.54% are longer than 50 residues (data not shown). Shorter motif length resulted in a similar number of allergen motifs with lower prediction accuracy (data not shown).

The log-odds matrix was extracted from the MEME output and converted into a generalized profile (19) with one match state for each position in the log-odds matrix. The profile was scaled on a randomized version of Swiss-Prot using pfscale (14).

The scaled profile was used to search allergens for matching sequences using a normalized score of 8.5 as threshold. This score corresponds to less than one chance match to be expected when searching whole Swiss-Prot and TrEMBL databases (roughly 700,000 sequences).

Matching allergens were removed from the allergen database, and remaining sequences were submitted to the next iteration of motif discovery.

Of 779 allergen sequences, 644 were matched by one or several of these motifs. Of the 135 sequences that did not match an allergen motif, 78 corresponded to partial allergen sequences and could therefore not be optimally aligned to an allergen motif, and the remaining 57 were assumed to represent relatively unique allergens. As we wanted the allergen motifs to represent the common characteristics of a group of related sequences, we decided not to generate potentially unrepresentative motifs for each of these 135 allergen sequences. Nevertheless, the 135 sequences were included in the allergenicity prediction (see below).

Motif-based allergenicity prediction

A query protein sequence was predicted to be allergenic, if it matched an allergen motif with a normalized score greater than 8.5. If no matching motif was found, a second analysis step was performed for allergenicity prediction: The query sequence was aligned to the allergen sequences not matching one of the allergen motifs (135 sequences). The query was rated allergenic if it either matched an allergen motif or scored better than a certain threshold (specified as a parameter) in a pair-wise sequence alignment. A length limit of 25 amino acids was applied. This corresponds to the length of the smallest known allergen protein (bee venom melittin, 26

residues, Swiss-Prot P01501). This length limit was applied in order to exclude short fragment sequences contained in sequence databases. Where indicated, the length limit was also applied for the FAO/WHO prediction method for better comparability of both methods. Sequence alignments were performed by using blastp from BLAST (17), BLOSUM50 (20) similarity matrix and 12/2 as gap open/extension penalties. BLAST E-values were as discriminative between true and false positives as PRSS (16) *P* values by using 3,000 reshufflings, and both performed better than raw scores or log length normalized raw scores (data not shown). For evaluation of prediction, a test database was constructed containing true allergen sequences and three randomized version of each allergen sequence (reversed, shuffled, 20 residue window-shuffled). The resulting test database contained 2,976 protein sequences and 25% true allergens. Performance of prediction was determined by calculating *precision* and *recall* (=sensitivity) defined as:

$$\text{precision} = \text{true}_{\text{positives}} / (\text{true}_{\text{positives}} + \text{false}_{\text{positives}})$$

$$\text{recall} = \text{true}_{\text{positives}} / (\text{true}_{\text{positives}} + \text{false}_{\text{negatives}})$$

Ten-fold cross validation experiment

The cross-validation experiment was performed by randomly splitting the allergen database into ten parts containing equal number of sequences. The sequences contained in each part were submitted to allergenicity prediction by both FAO/WHO and motif-based methods, whereas the remaining nine parts served as allergen reference database and as source for allergen motifs. Performance was measured as *precision* and *recall*, using the allergen sequences as true positives and three randomized versions of each allergen sequence as true negatives (reversed, shuffled, 20 amino acid window-shuffled).

RESULTS

Allergen sequence database

A database of allergen sequences that is as complete as possible represents a prerequisite for bioinformatic analysis of allergens, such as defining common allergen motifs or classifying new proteins according to their similarity with known allergens. Although most sequences of allergenic proteins are known and publicly available, no single database exists that contains all of these sequences (11).

Several lists of sequence accession numbers have been published corresponding to allergen genes or proteins (10, 11). Allergens contained in the Swiss-Prot protein database (6) are collected in a dedicated index. Thus, a complete allergen database was generated by extracting all accession numbers in the published allergen lists and downloading the corresponding sequences from the public sequence databases [Swiss-Prot (6), PIR (13) and GenBank (12)]. We have written a script automatically performing this task that allows frequent database updates and facilitates the error-prone and time-consuming process of downloading the sequences manually. The allergen database used in this study was generated on February 11, 2002, and

contained 779 non-redundant protein sequences, including translated allergen genes and generated sequence variants.

Evaluation of current allergenicity prediction

It is not known whether the current method for evaluation of allergenicity proposed by FAO/WHO has been tested concerning its recall and precision. We have therefore implemented the proposed method in a program and performed allergenicity prediction for a number of different databases as described in the experimental protocol. Table 1 shows the percentages of proteins predicted to be allergenic for 35% identical residues and different values of the parameter *n*. Using a value of 6 for the identity length *n* as proposed by FAO/WHO, 98.6% of the allergens in our database were correctly predicted. However, 67.3% of all proteins in Swiss-Prot were also rated as allergens, and this figure is reduced by only 0.08% if known allergens are removed from Swiss-Prot before analysis; 75.9% allergenic proteins were found in rice, and 42.9% of human trGEN sequences (7) (an automatically translated version of the human genome) were predicted to be allergenic.

For the prediction shown in Table 1, signal peptides were not removed from sequences as recommended by the FAO/WHO guidelines. The reason for this simplification was that only for a minority of analyzed proteins (5.6% of Swiss-Prot proteins, and none of the proteins in rice or trGEN databases), experimental evidence on the signal peptide was available in database annotation. We therefore did not truncate sequences for allergenicity prediction. Nevertheless, we studied whether cleaving of signal peptides might influence the prediction (*Table 1*, Swiss-Prot-SP). We found the numbers of predicted allergens slightly decreased if signal peptides were removed (from 67.3 to 66.3%).

Next, we investigated the influence of increasing identity length on predicted allergens in Swiss-Prot, rice, human trGEN and allergen databases (Table 1). Augmenting the value of *n* drastically reduced numbers of matching *n*-mers and thus put more importance on the similarity criterion (35% over 80 residues) of the prediction algorithm (data not shown). This resulted in a higher stringency obtained for prediction, even though the numbers of predicted allergens were still higher than the expected percentage of real allergens (~0.4% for Swiss-Prot based on Swiss-Prot allergen index). We therefore tried to find a new approach to quantify potential cross-reactivity of a query sequence with a known allergen.

Automated iterative motif discovery in allergen database

To assess variability contained in the allergen database and to generate a minimal set of sequence motifs representing allergens, an automated iterative motif discovery was performed. Only 52 statistically relevant allergen motifs were identified in the allergen database, indicating limited variability of allergen motifs in comparison to the total number of allergens contained in the database. Of 779 allergen sequences, 644 were matched by one or several of these motifs. Of the remaining sequences, 78 corresponded to short fragment allergen sequences that could therefore not be optimally aligned to an allergen motif. Thus, the 52 allergen motifs can match over 90% of allergens longer than 50 residues. Table 2 shows statistical motif qualities expressed as MEME E-values (15) for the 20 first-identified allergen motifs. E-values are an estimate of the

number of similar motifs to be expected by chance, with smaller values corresponding to more relevant motifs. If motif discovery was performed on a randomized version of our allergen database, the E-value of the best motif was $2.6 \cdot 10^{-12}$ (data not shown). The most frequent motif in the allergen database matched 101 proteins all belonging to the Bet v 1 family. This result can be explained by a bias in available allergen sequences toward well characterized birch pollen allergen Bet v 1 and related allergens, as well as by the high number of Bet v 1 isoforms. Der p 1, the clinically relevant major allergen from house dust mite, resides in the group of 16 proteins matching allergen motif 12 (AM00012). Four of the 20 allergen motifs shown in Table 2 (AM00004, AM00008, AM00017, and AM00020) could not be related to a known protein family. This finding emphasizes the necessity to use allergen-derived motifs for allergenicity prediction instead of using predefined protein family signatures as those in PROSITE or InterPro (21, 22).

Motif-based allergenicity prediction

The allergen motifs identified by iterative motif discovery were used to predict potential allergenicity of query protein sequences. The allergen motifs represent a collection of the sequence families present in currently known allergens. By scanning a query sequence with these motifs, its relatedness and thus its potential allergenicity can be estimated.

The 52 allergen motifs could not match the 135 sequences. Of these, 78 corresponded to partial allergen sequences and were significantly shorter than other allergens (data not shown). The remaining 57 sequences did not have closely related sequences in the allergen database and were therefore not represented in the allergen motif collection. To correctly predict also these sequences, we designed a two-step approach for allergenicity prediction. In the first step, query sequences are compared with allergen motifs. In the second step, query sequences are aligned to the 135 unique allergen sequences (not matching an allergen motif). A similarity identified in either of both steps indicates a potential cross-reactivity of the query sequence with a known allergen.

We first studied accuracy of the approach in a ten-fold cross validation experiment (Table 3). For this experiment, the allergen database has been split into 10 random parts of equal size. Allergenicity prediction was performed for sequences in each part, using the other nine parts as allergen reference database. This approach allowed estimation of prediction accuracy for so-far unknown allergens. Non-allergen sequences were generated by randomization of true allergens. In the cross-validation experiment, the FAO/WHO method (both for $n=6$ and $n=8$) proved more sensitive than the motif-based prediction (recall of 97.0% and 92.2% vs. 86.2%, Table 3). The high recalls obtained by both methods point out the limited variability of the allergen database; even if 10% of the sequences are removed, most can still be correctly classified as allergens. A wider divergence between allergenicity prediction methods was observed in measurements of precision. Whereas the motif-based method was highly accurate (precision of 94.8%, Table 3), the FAO/WHO method reached a precision of only 37.6%. Increasing the identity length parameter n of the FAO/WHO method from six to eight amino acids improved the precision to 68%. However, further increment of n did not result higher precision (data not shown).

Using a test database containing 2'976 protein sequences and 25% true allergens, we addressed the accuracy of allergenicity prediction methods. Non-allergen sequences in the test database have been generated by randomization of allergen sequences. Fig. 1 shows precision and recall of the motif-based allergenicity prediction and the prediction according to FAO/WHO guidelines by using various parameter values. Maximal precision and recall obtained by the motif-based prediction were superior to the ones obtained by the FAO/WHO method. Using a BLAST E-value cut-off of 10^{-8} (indicated by vertical line, Fig. 1A), the motif-based prediction reached a precision of 95.5% with a recall of 100%, whereas an identity length *n* of six amino acids for the FAO/WHO method (vertical line, Fig. 1B) resulted in a low precision of 36.6% with a recall of only 99.7%.

Finally, we directly compared motif-based and FAO/WHO prediction methods when applied to real protein sequences (Table 4). For all proteins contained in Swiss-Prot, allergenicity was predicted. As already shown in Table 1, more than two-thirds of the query proteins are predicted allergenic by the FAO/WHO method. Compared with this, motif-based prediction detects only 4% allergens in Swiss-Prot, and if allergen motifs are used exclusively for prediction (Table 4, motifs only method), this value is reduced further to 2.6%. To distinguish known allergens from false positives and potentially new allergens, we checked whether their sequence was contained in the allergen reference database. About 1 in 10 potential allergens predicted by the motif based methods was a true allergen, whereas only ~1 in 200 potential allergens was a true allergen when using the FAO/WHO method (Table 4).

DISCUSSION

Although the scientific community agrees on including sequence similarity in evaluation of allergenicity of foods derived from biotechnology (4, 5, 23, 24), no consensus has been reached on how to perform similarity testing (5). The aim of our study was to analyze allergen prediction methods on the basis of data acquired from known allergens and a large number of different proteins. To our knowledge, no such data-driven analysis has been performed so far. Considering the results we obtained for our reference allergen database and other general databases, we could quantify accuracy of allergen prediction. In addition, we could test a new approach for allergenicity prediction and could quantitatively compare it with current methods.

It must be pointed out that currently it cannot be claimed that a protein without sequence similarity to any known allergen might never cause an allergic reaction. Nevertheless, allergenicity prediction based on protein sequence provides an important tool to identify potential cross-reactivity with known allergens, indicating the requirement for further investigation by other techniques (4, 25).

Allergen sequence database

As allergens do not share common structural characteristics (26, 27), and epitopes recognized by the immune system cannot be predicted based on sequence data, the use of sequence similarity in allergenicity evaluation is highly dependent on a database of allergens that serves as reference. Instead of manually constructing an allergen database by literature review and database searching, we relied on previously published and regularly updated allergen lists (3, 6, 10). Our

focus was to obtain a database that would be as comprehensive as possible and to overcome the shortcomings of currently existing databases (11). Special attention was paid to sequence variants: In Swiss-Prot, variants are not contained in the database as separate sequences, but only as annotation accompanying the principal sequence entry. Processing of variant information yielded an additional 99 sequences that would otherwise not have been included in the allergen database. It will be important to update the allergen reference database on a regular basis as new allergens are identified in order to improve the performance of allergenicity prediction.

Evaluation of current allergenicity prediction

Based on current knowledge, it is not justified to consider each protein with six contiguous amino acids in common with a known allergen as potentially allergenic. This criterion predicts the majority of Swiss-Prot or rice proteins and more than 40% of human proteins as allergens, which does not reflect the numbers of true allergens to be expected in these databases. In addition, the numbers of matching 6-mers found in a protein sequence tend to increase with sequence length (data not shown). This is indicative of this prediction method producing mainly random matches whose probability increases with sequences length. FAO/WHO guidelines recommend to remove signal peptides from allergen and query sequences before allergenicity prediction. The resulting decrease of allergens predicted in Swiss-Prot may be explained by the simultaneous decrease of sequence length, and hence, a lowered probability of matching 6-mers. Prediction accuracy was not affected by signal peptide removal (data not shown).

We could show that by increasing the identity length parameter, the precision of FAO/WHO allergenicity prediction could be improved without affecting its sensitivity. However, values larger than 8 did not result in further performance gains. An adjustment of the similarity parameters (35% over 80 residues) might be necessary to optimize the performance of the approach, especially as the current values are chosen conservatively: Allergenic cross-reactivity caused by proteins sharing conformational or linear epitopes is rare at 50% identity and typically requires more than 70% amino acid identity across the full length of the proteins (26). However, we think that a motif based method is superior to a conventional sequence alignment method, as it is more flexible (see below).

Allergen motifs identified from allergen database

An improved prediction performance can thus be obtained by increasing the identity length from 6 to 8 residues and by optimizing FASTA or BLAST alignment parameters. However, local alignment search tools, such as FASTA and BLAST, exert fixed substitution scores and gap penalties, and one single similarity cut-off value would have to be defined that could discriminate between immunologically cross-reactive and non-cross-reactive proteins. Indeed, such a universal cut-off value may not exist, and individual thresholds may be necessary for different protein families. Hence, we choose to detect common sequence motifs by using profiles for allergenicity prediction that would provide us with the necessary flexibility. Profiles such as those used in our study (19) are more sensitive in detecting homologues and thus potentially cross-reactive proteins than local alignment search tools, because of their position-specific scoring system (28). In addition, each individual profile representing a motif was scaled (14)

such as the match scores produced by motif searching all become normalized and thus comparable. This is the basis for a universal threshold for immunological cross-reactivity.

Furthermore, allergen motifs serve to systematically organize allergens into groups of related and cross-reactive proteins. Our data indicate that amongst currently known allergens, there are 52 sequence families that are represented by more than one allergenic protein, and an additional group of 135 fragment sequences or unique allergens without allergenic relatives. The allergen motifs we used were identified and scaled according to an automated protocol. We have shown that allergenicity prediction based on these motifs is possible with high sensitivity and greatly improved precision compared with the current method. Nevertheless, manual inspection, such as realigning motif-containing allergens and construction of optimized profiles, has the potential to further increase prediction performance. It would be possible to include spatial information from three-dimensional structure models into the profiles, for instance to focus the profile on core residues that define the overall protein fold, and on surface-accessible residues that may be essentially determining the characteristic properties of IgE binding epitopes. Such improvements could not be realized when using local alignment search tools, as it has been proposed by FAO/WHO and by others (23). In the future, increasing numbers of allergenic proteins will be identified, resulting in a more complete set of allergen motifs and probably eliminating the need to perform pairwise alignments in motif based allergenicity prediction.

Motif-based allergenicity prediction

In a 10-fold cross validation experiment, we addressed the performance of allergenicity prediction for new allergens not contained in the allergen reference database. Both prediction methods are evidently highly sensitive, although the high recall attained by prediction according to FAO/WHO guidelines has to be ascribed to the high level of false positives produced by the method.

Precision of prediction methods was assessed by using a test set containing 25% true allergens. The non-allergenic sequences in the test set were obtained by randomization of true allergen sequences. This procedure was chosen because it alters protein fold and thus immunological properties of randomized sequences, whereas it preserves other sequence characteristics, such as compositional bias and low complexity regions, that are known to

produce statistically relevant but biologically meaningless matches (29). The results obtained for this test set (Fig. 1) are consistent with our earlier findings, namely that both methods are highly sensitive, but the FAO/WHO method produces a high number of false positives, reducing its precision. This is especially true for parameters as proposed by the FAO/WHO.

Finally, we wanted to test the prediction methods on real protein sequences. Allergenicity prediction was performed for all proteins contained in Swiss-Prot (Table 4). Only 1 in 200 potential allergens predicted according to FAO/WHO guidelines was a true allergen. It is evident that with such a high level of noise, the method cannot discriminate between non-allergens and allergens. Moreover, the method could lead to a general overestimation of allergenic potential and thus require disproportionately high efforts in clinical risk assessment. One in 12 allergens detected by the motif based method was a true allergen, and if allergenicity prediction was

performed by exclusively matching allergen motifs, 1 in 9 detected sequences was a true allergen. Indeed, the percentages of true allergens predicted by allergen motifs might be underestimated, as some of the potential proteins could actually be cross-reactive. Because of its enhanced signal-to-noise ratio, this method could be used to systematically search for new potential allergens in available sequence data.

Unfortunately, it is currently not possible to define a similarity threshold in allergenicity prediction that can truly discriminate between immunologically cross-reactive and non-cross-reactive proteins. More information on the relationship between sequence and structure and eventually adapted similarity search tools are needed for such a future estimate. Manually optimized profiles that imply information from a known three-dimensional structure may be a promising option.

Although the properties conferring allergenicity remain unknown, we have shown that our approach predicts allergenicity with good sensitivity and precision, and it performs definitely better than the current method proposed by the FAO/WHO. With growing allergen catalogs and improved methods for characterizing proteins, our approach may be further optimized and provides a reasonable tool in risk assessment to identify transgenes that require further investigation by other techniques.

ACKNOWLEDGMENTS

This work was supported by the Novartis Foundation, grant 02A07.

REFERENCES

1. Nadler, M. J., Matthews, S. A., Turner, H., and Kinet, J. P. (2000) Signal transduction by the high-affinity immunoglobulin E receptor Fc epsilon RI: coupling form to function. *Adv. Immunol.* **76,** 325–355

2. Metzger, H. (1991) The high affinity receptor for IgE on mast cells. *Clin. Exp. Allergy* **21,** 269–279

3. Gendel, S. M. (2002) Sequence analysis for assessing potential allergenicity. *Ann. N. Y. Acad. Sci.* **964,** 87–98

4. Metcalfe, D. D., Astwood, J. D., Townsend, R., Sampson, H. A., Taylor, S. L., and Fuchs, R. L. (1996) Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Crit. Rev. Food Sci. Nutr.* **36, Suppl,** S165–S186

5. Taylor, S. L., and Hefle, S. L. (2001) Will genetically modified foods be allergenic? *J. Allergy Clin. Immunol.* **107,** 765–771

6. Bairoch, A., and Apweiler, R. (1996) The SWISS-PROT protein sequence data bank and its new supplement TREMBL. *Nucleic Acids Res.* **24,** 21–25

7. Pagni, M., Iseli, C., Junier, T., Falquet, L., Jongeneel, V., and Bucher, P. (2001) trEST, trGEN and Hits: access to databases of predicted protein sequences. *Nucleic Acids Res.* **29,** 148–151

8. Quackenbush, J., Liang, F., Holt, I., Pertea, G., and Upton, J. (2000) The TIGR gene indices: reconstruction and representation of expressed gene sequences. *Nucleic Acids Res.* **28,** 141–145

9. Quackenbush, J., Cho, J., Lee, D., Liang, F., Holt, I., Karamycheva, S., Parvizi, B., Pertea, G., Sultana, R., and White, J. (2001) The TIGR Gene Indices: analysis of gene transcript sequences in highly sampled eukaryotic species. *Nucleic Acids Res.* **29,** 159–164

10. King, T. P., Hoffman, D., Lowenstein, H., Marsh, D. G., Platts-Mills, T. A., and Thomas, W. (1994) Allergen nomenclature. WHO/IUIS Allergen Nomenclature Subcommittee. *Int. Arch. Allergy Immunol.* **105,** 224–233

11. Gendel, S. M. (1998) Sequence databases for assessing the potential allergenicity of proteins used in transgenic foods. *Adv. Food Nutr. Res.* **42,** 63–92

12. Benton, D. (1990) Recent changes in the GenBank On-line Service. *Nucleic Acids Res.* **18,** 1517–1520

13. George, D. G., Barker, W. C., Mewes, H. W., Pfeiffer, F., and Tsugita, A. (1996) The PIR-International Protein Sequence Database. *Nucleic Acids Res.* **24,** 17–20

14. Bucher, P., Karplus, K., Moeri, N., and Hofmann, K. (1996) A flexible motif search technique based on generalized profiles. *Comput. Chem.* **20,** 3–23

15. Bailey, T. L. and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28–36

16. Pearson, W. R. (2000) Flexible sequence similarity searching with the FASTA3 program package. *Methods Mol. Biol.* **132,** 185–219

17. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402

18. Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M., and Sonnhammer, E. L. (2002) The Pfam protein families database. *Nucleic Acids Res.* **30,** 276–280

19. Bucher, P. and Bairoch, A. (1994) A generalized profile syntax for biomolecular sequence motifs and its function in automatic sequence interpretation. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 53–61

20. Henikoff, S., and Henikoff, J. G. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA*. **89,** 10915–10919

21. Falquet, L., Pagni, M., Bucher, P., Hulo, N., Sigrist, C. J., Hofmann, K., and Bairoch, A. (2002) The PROSITE database, its status in 2002. *Nucleic Acids Res.* **30,** 235–238

22. Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M. D., et al. (2001) The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res.* **29,** 37–40

23. Hileman, R. E., Silvanovich, A., Goodman, R. E., Rice, E. A., Holleschak, G., Astwood, J. D., and Hefle, S. L. (2002) Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. *Int. Arch. Allergy Immunol.* **128,** 280–291

24. Gendel, S. M. (1998) The use of amino acid sequence alignments to assess potential allergenicity of proteins used in genetically modified foods. *Adv. Food Nutr. Res.* **42,** 45–62

25. Astwood, J. D., and Fuchs, R. L. (1996) Allergenicity of foods derived from transgenic plants. *Monogr. Allergy* **32,** 105–120

26. Aalberse, R. C. (2000) Structural biology of allergens. *J. Allergy Clin. Immunol.* **106,** 228–238

27. Bredehorst, R., and David, K. (2001) What establishes a protein as an allergen? *J. Chromatogr. B Biomed. Sci. Appl.* **756,** 33–40

28. Gribskov, M., Luthy, R., and Eisenberg, D. (1990) Profile analysis. *Methods Enzymol.* **183,** 146–159

29. Wootton, J. C., and Federhen, S. (1996) Analysis of compositionally biased regions in sequence databases. *Methods Enzymol.* **266,** 554–571

Received November 26, 2002; accepted February 27, 2003.

Table 1

Allergen prediction according to FAO/WHO guidelines.

Database Version (number of proteins)	*% allergens for a given identity length n*[a]						
	n=6	*n=7*	*n=8*	*n=9*	*n=10*	*n=11*	*n=12*
Allergens 02-11-2002 (779)	98.6	98.2	98.0	97.7	97.6	97.3	97.0
Swiss-Prot Release 40.0 (101'602)	67.3	17.6	8.7	7.6	7.3	7.2	7.2
Swiss-Prot w/o Allergens[b] Release 40.0 (101'328)	67.3	17.4	8.5	7.3	7.0	7.0	7.0
Swiss-Prot-SP[c] Release 40.0 (101'602)	66.3	17.1	8.6	7.5	7.2	7.2	7.2
Rice TIGR OsGI Release 7.0 (10'891)	75.9	27.6	11.3	8.0	7.3	7.2	7.2
trGEN human 12-19-2001 (330'743)	42.9	7.3	2.9	2.2	2.1	2.0	2.0

[a] Query proteins were rated allergenic if either at least *n* consecutive residues were found in common with a known allergen, or if sequence identity with a known allergen was higher than 35 % over a window of 80 residues.

[b] 274 proteins listed in Swiss-Prot allergen index were removed from Swiss-Prot.

[c] Swiss-Prot sequences with known signal peptides were truncated according to annotation (FT SIGNAL) and stored in the database termed Swiss-Prot-SP.

Table 2

Allergen motifs.

Motif identifier	*MEME E-value*	*Matching allergens*	*Predominant protein families*[a]
AM00001	$1.8 \cdot 10^{-4123}$	101	Pathogenesis-related proteins BetvI family
AM00002	$2.0 \cdot 10^{-1477}$	68	Profilins Pollen proteins Ole e I family
AM00003	$1.3 \cdot 10^{-919}$	36	Globins
AM00004	$3.0 \cdot 10^{-845}$	35	-[b]
AM00005	$4.8 \cdot 10^{-794}$	22	SCP/Tpx-1/Ag5/PR-1/Sc7
AM00006	$2.3 \cdot 10^{-774}$	34	11-S plant seed storage proteins Caseins
AM00007	$9.2 \cdot 10^{-460}$	47	Plant lipid transfer proteins Lipases Eukaryotic thiol proteases
AM00008	$2.7 \cdot 10^{-460}$	18	-[b]
AM00009	$1.9 \cdot 10^{-323}$	14	EF-hand calcium-binding domain
AM00010	$3.5 \cdot 10^{-271}$	10	Cereal trypsin/alpha-amylase inhibitors
AM00011	$2.0 \cdot 10^{-356}$	11	Tropomyosins
AM00012	$3.2 \cdot 10^{-242}$	16	Eukaryotic thiol proteases
AM00013	$1.2 \cdot 10^{-234}$	8	Mitochondrial energy transfer proteins
AM00014	$1.6 \cdot 10^{-229}$	7	Lipocalins
AM00015	$3.3 \cdot 10^{-219}$	21	Uteroglobin family Serpins
AM00016	$4.1 \cdot 10^{-218}$	9	Caseins
AM00017	$1.2 \cdot 10^{-190}$	7	-[b]
AM00018	$7.8 \cdot 10^{-165}$	24	Plant lipid transfer proteins Chitin binding domain Barwin domain
AM00019	$2.1 \cdot 10^{-139}$	6	Enolases
AM00020	$1.3 \cdot 10^{-211}$	12	-[b]

[a]Predominant protein families corresponding to allergen motifs have been identified by scanning motif containing sequences with PROSITE (21), Rel. 16.0 and updates up to Oct 2001.

[b] No predominant PROSITE protein family signature was found in matching allergens.

Table 3

Ten-fold cross validation.

Dataset[a]	*Allergen sequences*	*Motifs*	*% precision*			*% recall*		
			FAO/ WHO (n=6)	*FAO/ WHO (n=8)*	*motif based method*	*FAO/ WHO (n=6)*	*FAO/ WHO (n=8)*	*motif based method*
Set 0	75	50	36.9	68.6	97.0	97.3	93.3	86.7
Set 1	76	53	36.4	62.0	90.5	98.7	92.1	88.2
Set 2	75	49	40.0	72.5	96.9	98.7	94.7	84.0
Set 3	74	51	38.1	63.9	94.0	97.3	93.2	85.1
Set 4	77	46	37.5	67.3	98.6	97.4	93.5	88.3
Set 5	74	51	37.8	69.6	97.1	96.0	96.0	90.5
Set 6	73	49	39.4	76.2	96.8	97.3	87.7	82.2
Set 7	75	52	36.7	68.4	95.4	93.3	86.7	82.7
Set 8	70	44	36.3	66.0	87.0	98.6	94.3	85.7
Set 9	75	46	36.7	68.0	95.7	96.0	90.7	88.0
TOTAL	744	-	37.6	68.0	94.8	97.0	92.2	86.2

[a]Three randomized versions of each allergen sequence in the set were generated (reversed, shuffled, 20 amino acid window-shuffled). The allergen sequences (true positives) and the randomized sequences (true negatives) were submitted to allergenicity prediction, using all other datasets as allergen reference database and a lower length limit of 25 residues.

Table 4

Comparison of FAO/WHO and motif based allergen predictions for Swiss-Prot proteins.

Prediction method	*Potential allergens*[a]		*True allergens*[b]	
FAO/WHO	68356	67.3 %	351	0.5 %
Motif based	4093	4.0 %	351	8.6 %
Motifs only[c]	2603	2.6 %	297	11.4 %

[a]Predicted allergens for Swiss-Prot proteins longer than 25 residues.
[b]A potential allergen was considered a true allergen, if its sequence was contained in the allergen reference database.
[c]For the "motifs only" method, only allergen motifs were used for allergenicity prediction, without local similarity search as in step two of motif-based method.

Fig. 1

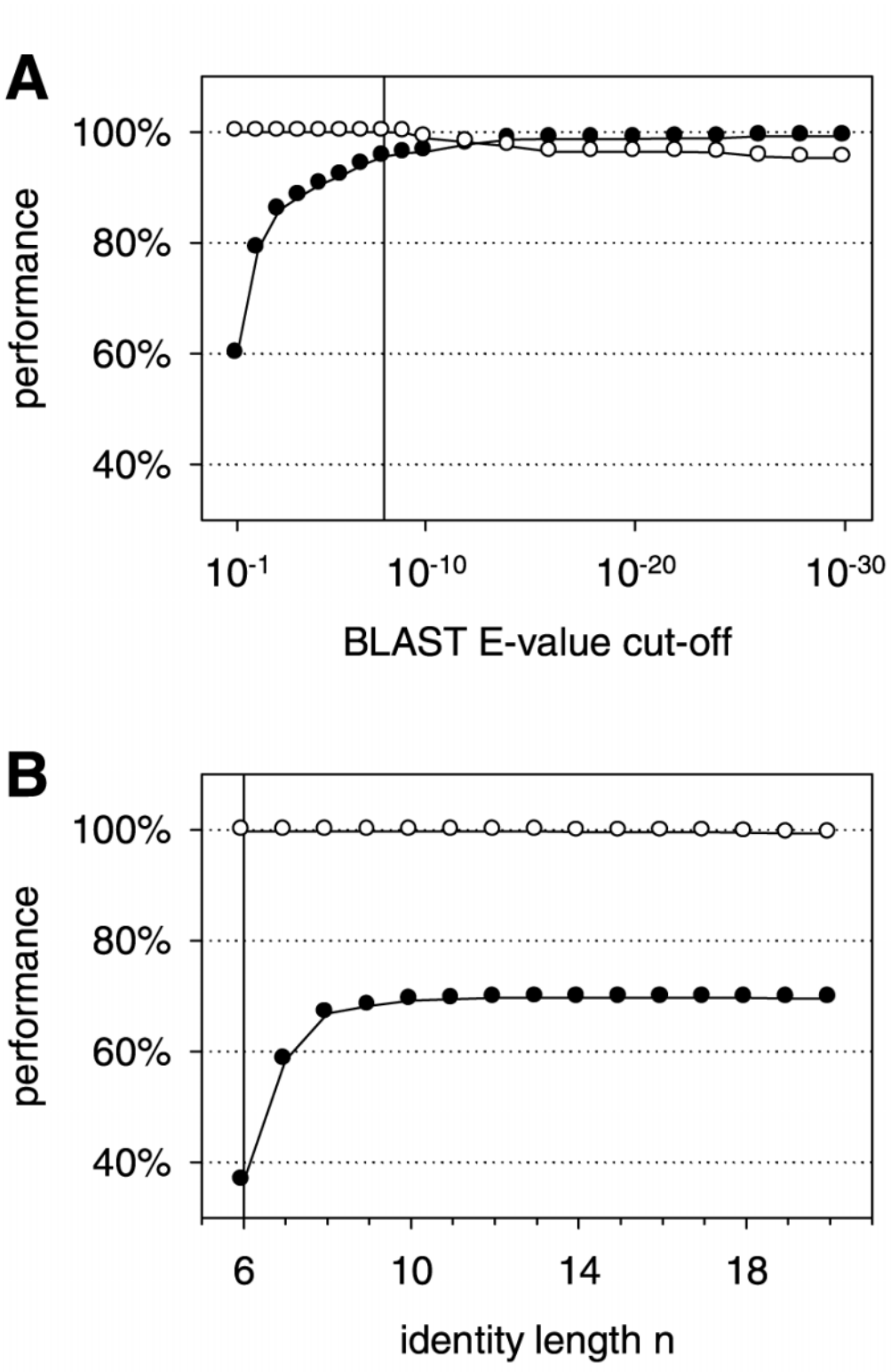


Figure 1. Performance of motif based allergenicity prediction and of prediction according to FAO/WHO guidelines. A test database containing 2,976 protein sequences and 25 % true allergens has been submitted to allergenicity prediction by motif based prediction strategy (**A**) and according to FAO/WHO guidelines (**B**), using a lower length limit of 25 residues for both methods. Prediction stringency has been varied by BLAST E-value cut-off (used in second analysis step of motif based prediction) and identity length for FAO/WHO prediction method. Prediction performance is shown as % precision (●) and % recall (○). The vertical lines indicate BLAST E-value cut-off 10^{-8} and identity length n=6.

Appl Microbiol Biotechnol (2002) 59:426–435
DOI 10.1007/s00253-002-1032-6

MINI-REVIEW

E. Schuster · N. Dunn-Coleman · J. C. Frisvad
P. W. M. van Dijck

On the safety of *Aspergillus niger* – a review

Received: 6 March 2002 / Revised: 17 April 2002 / Accepted: 19 April 2002 / Published online: 25 June 2002

Abstract *Aspergillus niger* is one of the most important microorganisms used in biotechnology. It has been in use already for many decades to produce extracellular (food) enzymes and citric acid. In fact, citric acid and many *A. niger* enzymes are considered GRAS by the United States Food and Drug Administration. In addition, *A. niger* is used for biotransformations and waste treatment. In the last two decades, *A. niger* has been developed as an important transformation host to over-express food enzymes. Being pre-dated by older names, the name *A. niger* has been conserved for economical and information retrieval reasons and there is a taxonomical consensus based on molecular data that the only other common species closely related to *A. niger* in the *Aspergillus* series Nigri is *A. tubingensis*. *A. niger*, like other filamentous fungi, should be treated carefully to avoid the formation of spore dust. However, compared with other filamentous fungi, it does not stand out as a particular problem concerning allergy or mycopathology. A few medical cases, e.g. lung infections, have been reported, but always in severely immunocompromised patients. In tropical areas, ear infections (otomycosis) do occur due to *A. niger* invasion of the outer ear canal but this may be caused by mechanical damage of the skin barrier. *A. niger* strains produce a series of secondary metabolites, but it is only ochratoxin A that can be regarded as a mycotoxin in the strict sense of the word. Only 3–10% of the strains examined for ochratoxin A production have tested positive under favourable conditions. New and unknown isolates should be checked for ochratoxin A production before they are developed as production organisms. It is concluded, with these restrictions, that *A. niger* is a safe production organism.

Introduction

Aspergillus niger has been the subject of research and industrial use for several decades. It first acquired practical importance in 1919, when its ability to produce citric acid was industrially exploited. Gluconic and fumaric acids have been produced with *A. niger*, although they are of less economic importance. However, since the 1960s, *A. niger* has become a source of a variety of enzymes that are well established as technical aids in fruit processing, baking, and in the starch and food industries. Gene technology has been successfully applied to improve production processes and to make use of *A. niger* as an expression system for foreign proteins. The intense research over the past decade has resulted in a range of new processes and products.

Ecology

Many black *Aspergilli* have been isolated from all over the world. *A. niger* is a filamentous fungus growing aerobically on organic matter. In nature, it is found in soil and litter, in compost and on decaying plant material. Reiss (1986) collected data on the influence of temperature, water activity and pH on the growth of various

Dr. Schuster (presently retired from Röhm GmbH – Röhm is now AB Enzymes GmbH) prepared the first draft of this manuscript in 1993 on behalf of the AD HOC 5 working group of AMFEP [Association of Manufacturers of Fermentation Enzyme Products, Brussels]; later on, contributions from Dr. Dunn-Coleman and Dr. Frisvad were included. The present manuscript is a revised and updated version prepared by Drs. Frisvad and Van Dijck.

E. Schuster
Enzyme Technology, Röhm GmbH, Kirschenallee 45,
64293 Darmstadt, Germany

N. Dunn-Coleman
Genencor International Inc., 925 Page Mill Road,
Palo Alto, CA 94304, USA

J.C. Frisvad
BioCentrum-DTU, Building 221,
Technical University of Denmark,
2800 Kgs. Lyngby, Denmark

P.W.M. van Dijck (✉)
DSM Food Specialties, Department of Regulatory Affairs,
600–0245, P.O. Box 1, 2600MA, Delft, The Netherlands
e-mail: piet.dijck-van@dsm.com

Aspergilli. *A. niger* is able to grow in the wide temperature range of 6–47°C with a relatively high temperature optimum at 35–37°C. The water activity limit for growth is 0.88, which is relatively high compared with other *Aspergillus* species. *A. niger* is able to grow over an extremely wide pH range: 1.4–9.8. These abilities and the profuse production of conidiospores, which are distributed via the air, secure the ubiquitous occurrence of the species, with a higher frequency in warm and humid places (Rippel-Baldes 1955).

Taxonomy

Raper and Fennel (1965) divided the genus *Aspergillus* into groups according to the colour of the conidiospores. *Aspergilli* with brown to black-shaded spores constitute the *A. niger* group. Although the members of this group vary considerably, only a few differ so clearly from the majority that they can easily be classified as separate species (e.g. *A. carbonarius*, *A. japonicus*, *A. ellipticus*, *A. heteromorphus* and *A. aculeatus*). Most of the brown to black *Aspergilli* belong to the other group of species, which are difficult to distinguish: *A. ficuum*, *A. phoenicis*, *A. niger* and *A. awamori* being the most prominent. In practice, this group of species is often together called *A. niger* van Tieghem.

The apparently insignificant differences between members of the *A. niger* group were the decisive reasons for Al-Musallam (1980) to classify some species as varieties of *A. niger*, while Raper and Fennell considered them to be separate species.

Introducing restriction fragment length polymorphism (RFLP) analysis to *Aspergillus* taxonomy, Kusters-van Someren et al. (1990, 1991) analysed the ribosomal banding patterns and the hybridisation patterns of genomic digests from strains in the *A. niger* group, using pectin lyase genes as probes for hybridisation. They proposed a different, but in their opinion more reliable, classification of the *A. niger* group, reducing the number of species from 13 (Raper and Fennel 1965) to 6: *A. carbonarius*, *A. japonicus*, *A. ellipticus*, *A. heteromorphus*, *A. niger* and *A. tubingensis*, the latter of which consists of strains formerly called *A. niger*. The difference between *A. niger* and *A. tubingensis* has been repeatedly confirmed in further studies using RFLP of mitochondrial DNAs and ribosomal repeat units (Varga et al. 1993, 1994; Parenicová et al. 1997), but also in studies using randomly amplified polymorphic DNA (Megnegneau et al. 1993); internal transcribed spacer sequence data (Accensi et al. 1999; Parenicová et al. 2001), and nuclear genes encoding polygalacturonases, arabinoxylan-arabinofuranohydrolase and xylanases (Bussink et al. 1991; de Graaf et al. 1994; Gielkens et al. 1997). Despite this agreement on the molecular separation of *A. niger* and *A. tubingensis*, no phenotypic differences have yet been found between the two species (Varga et al. 2000; Parenicová et al. 2000). This may be the reason that *A. tubingensis* is not yet listed among species of *Aspergillus* in current use (Pitt and Samson 2000; Pitt et al. 2000).

The name *A. niger* is predated by the names *A. phoenicis* and *A. ficuum* and, provided it is accepted that these three taxa are all conspecific, as implied by most molecular studies (Parenicová et al. 2000; Varga et al. 2000), the latter two taxa would have nomenclatural priority. Since these latter two names are nowadays rarely used, it was proposed at the Second International Workshop on *Penicillium* and *Aspergillus* that the name *A. niger* was to be conserved and *A. phoenicis* and *A. ficuum* to be rejected (Frisvad et al. 1990; Kozakiewicz et al. 1992). *A. niger* is a species of major economic importance and the name is now conserved for practical, information retrieval and economical reasons, and in the interest of continuity in legal affairs and approval procedures. Despite this, *A. phoenicis*, *A. awamori* and *A. foetidus* are still mentioned in the list of types of species in current use (Pitt and Samson 2000), whereas the species *A. tubingensis*, *A. acidus*, and *A. citricus* are not listed (Kozakiewicz 1989). A case can be made for keeping the name *A. awamori* for the domesticated form of *A. niger* in parallel with keeping the name *A. oryzae* for the domesticated form of *A. flavus*, *A. sojae* for the domesticated form of *A. parasiticus*, and *Penicillium camemberti* for the domesticated form of *P. commune* (Pitt et al. 2000). The appearance on this list of *A. phoenicis* and *A. foetidus* is dubious, however.

Today's practice, that the designation *A. niger* van Tieghem includes strains which could be named *A. awamori*, *A. ficuum*, *A. foetidus*, *A. phoenicis*, *A. pulverulentus*, *A. tubingensis*, *A. inuii* and *A. usamii* should be continued. If, however, phenotypic differences are found between *A. niger* and *A. tubingensis* this may lead to acceptance of the latter taxon. Some molecular data have also indicated that *A. foetidus* (Parenicová et al. 2000) and a *nomen nudum A. brasiliensis* (Varga et al. 2000) are distinct species, but phenotypic data backing this up are meagre at best. The species currently acceptable in section *Nigri* are listed in Table 1.

The most recent supraspecific scheme for the genus *Aspergillus* was suggested by Gams et al. (1985), placing all species with dark brown to black-shaded conidia into the section *Nigri* of a proposed subgenus *Circumdati*. Their proposal has been accepted by the International Commission on *Penicillium* and *Aspergillus* (Samson 1992).

Industrial use

A. niger became an industrially used organism when citric acid was first produced by fermentation in 1919. Citric acid is widely used in a variety of industries and, by sales volume, greatly exceeds other metabolites such as gluconic acid (Roukas 2000). Citric acid is the primary acidulant in the food and beverage industries. It is used in foods such as soft drinks, fruit juices, desserts, jams, jellies, candy and wine. In the pharmaceutical in-

Table 1 Currently accepted species in *Aspergillus* subgenus *Circumdati* section *Nigri* (Kusters-van Someren et al. 1991; Parenicová et al. 2000, 2001)

Species generally accepted	Ochratoxin A production potential
A. niger	+ (low frequency)
A. carbonarius	++ (high frequency)
A. ellipticus	–
A. heteromorphus	–
A. aculeatus	–
A. japonicus	–
Species distinguishable with molecular data only:	
A. tubingensis (= *A. acidus* = *A. acidus* var. *pallidus*)	–
Species accepted by some authors (but = *A. niger* based on molecular data)	
A. foetidus	–
A. citricus	++ (high frequency)

dustry, iron citrate is used as a source of iron and citric acid as a preservative for stored blood; in the cosmetics and toiletries industries it is used as a buffer, for pH adjustment and as an anti-oxidant. It is also used in industrial applications including detergents, leather tanning, in electroplating and other applications where sequestering agent activity in the neutral to low pH range is required. Citric acid is produced almost exclusively by fermentation of *A. niger* and *A. wentii* because yields of these organisms are economic and formation of undesired side products is minimal. The Food and Drug Administration (FDA) has listed *A. niger* as a source of citric acid (21 Code of Federal Regulations §173.280).

In addition to citric acid, *A. niger* is a rich source of enzymes. Pectinase, protease and amyloglucosidase were the first to be exploited, and were originally produced in surface culture (Frost and Moss 1987). Although it had been shown by Kluyver's group in Delft as early as 1932 that it was possible to cultivate a filamentous fungus like *A. niger* in submerged culture (Kluyver and Perquin 1932), the technology was first applied to the production process of penicillin G by *Penicillium chrysogenum* in 1942. After 1950, production technology for fungal products gradually changed from surface culture to stirred-tank processes, but up until the mid 1960s companies used surface culture processes (Barbesgaard et al. 1992). Several additional enzymes like cellulase and hemicellulase were manufactured using black *Aspergillus* strains in stirred tank processes.

For the manufacture of many products, starch – one of the most abundant carbohydrates – must be hydrolyzed to syrups, which contain glucose, maltose and low molecular weight dextrins. Amyloglucosidase, also referred to as glucoamylase, is an exo-amylase catalysing the release of successive glucose units from the non-reducing ends of starch by hydrolysing α-1,4-D-glucosidic linkages. The glucose syrup and the alcohol industries are the principal users of amyloglucosidase produced by *A. niger*.

Pectin, a heteropolysaccharide, is a principal component in commercially important fruits and vegetables. Several enzymes, including pectin esterases, endo- and exo-polygalacturonidases and pectin lyases, produced from *A. niger* degrade pectin; they are used in wine and fruit juice production to reduce juice viscosity before pressing and improve clarification (Grassin and Fauguenbergue 1999).

It is established practice to improve the baking process by adding hemicellulases from *A. niger* when mixing the dough. The enzymes modify the rheological properties of the dough and give higher loaf volume and better crumb structure of bread and pastry.

A. niger glucose oxidase and catalase are used for determination of glucose mainly in diagnostic enzyme kits, for the removal of either glucose or oxygen from foods and beverages and for the production of gluconic acid from glucose (Berka et al. 1992).

FAO/WHO experts have repeatedly reviewed and accepted enzyme preparations from *A. niger* including the organism itself (FAO/WHO 1972, 1978, 1981, 1987, 1990), listing them with an Acceptable Daily Intake of 'not specified'. The FDA in the United States has accepted numerous enzymes for food use: in the early 1960s the FDA issued opinion letters recognizing that α-amylase, cellulase, amyloglucosidase, catalase, glucose oxidase, lipase and pectinase from *A. niger* can be 'generally regarded as safe' (GRAS) under the condition that non-pathogenic and non-toxigenic strains and current good manufacturing practices be used in production. In addition to these enzymes, Godfrey and Reichelt (1983) claimed GRAS status for β-galactosidase and protease from *A. niger*. Carbohydrase and cellulase from *A. niger* are also approved as a secondary direct food additive by the FDA as an aid in clam and shrimp processing (21 Code of Federal Regulations §173.120).

Until the 1980s, *A. niger* industrial production strains were isolated through the use of classical mutagenesis followed by screening and/or selection. Parasexual crossing has also been used in strain improvement efforts in *Aspergilli*, which lack a sexual cycle. For example, Das and Roy (1978) have reported improved production of citric acid by a diploid strain of *A. niger* generated by parasexual crossing.

With the development of DNA-mediated transformation of *Aspergilli*, initially in *A. nidulans* (Ballance et al. 1983; Tilburn et al. 1983), and subsequently in *A. niger*

(Buxton et al. 1985; Kelly and Hynes 1985; Van Hartingsveldt et al. 1987; Ward et al. 1988; Campbell et al. 1989), this very useful technology was applied to using *A. niger* as a host for gene expression. For example, the production of native *A. niger* catalase has been increased using recombinant techniques (Berka et al. 1994a, b), whereas a 1,000-fold improvement in the expression level for *A. niger* phytase was achieved by using recombinant technology (Van Gorcom et al. 1991; Van Hartingsveldt et al. 1993; Selten 1994).

The long history of safe use on an industrial scale makes *A. niger* exceptionally well suited to be used as a host for heterologous expression. A number of genes of commercial importance and their regulatory sequences that could be used as components in industrial expression systems have been cloned (Nunberg et al. 1983, 1984; Bussink et al. 1990; Harmsen et al. 1990; Nguyen et al. 1991).

The strategy of employing the promoter of a highly expressed fungal gene for the expression of a heterologous gene (Cullen et al. 1987) proved successful when Dunn-Coleman et al. (1991) obtained expression by *A. niger* var. *awamori* of commercially viable levels of calf chymosin under the control of the glucoamylase promoter. This enzyme has been accorded GRAS status by the FDA (Federal Register 1993).

After the cloning of the phytase gene (Mullaney et al. 1991; Van Gorcom et al. 1991; Van Hartingsveldt et al. 1993; Piddington et al. 1993) the gene cloned from an *A. niger* strain was inserted in an expression cassette under the control of the strong glucoamylase promoter. This expression cassette was randomly integrated in multiple copies in the genome of an industrial *A. niger* glucoamylase production strain (Selten 1994; Van Dijck 1999). One of the reasons for the high production level of glucoamylase in this particular strain is the multiplication of a region in the DNA containing, among other things, the promoter and coding sequence of the glucoamylase gene, the *gla*A-locus. Using advanced proprietary genetic modification techniques (Selten et al. 1995, 1998), this locus was "emptied" and subsequently "filled" with "genes of interest" in expression cassettes under the control of the host *gla*-promoter. Compared to the original glucoamylase overproducing strain they are completely identical except for the fact that the "gene of interest" replaces the glucoamylase gene (Groot et al. 2000). This has several advantages. New production strains can now be designed and built in a predictable manner. In addition, and this is important from a regulatory point of view, this technique of targeted integration by definition cannot cause any pleiotropic efforts by perturbing the rest of the genome. This is often raised as a (hypothetical) possibility by regulatory bodies in case of product approvals where the production strain was obtained by random integration of genes.

To date, many enzyme products are available on the market from recombinant strains of *A. niger.* In a recent listing prepared by the Enzyme Technical Association, the enzymes α-amylase, arabinofuranosidase, catalase, chymosin, glucoamylase, glucose oxidase, pectin esterase, phospholipase A2, phytase, and xylanase are mentioned as being produced by recombinant strains of *A. niger* (Pariza and Johnson 2001).

Safety aspects

A. niger is generally regarded as a safe organism. This is documented in lists of the organisations responsible for occupational health and safety [e.g. Berufsgenossenschaft der Chemischen Industrie (1998)]. In rare cases when persons are exposed to intense spore dust, hypersensitivity reactions have been observed.

Pathogenicity

A. niger is generally regarded as a non-pathogenic fungus widely distributed in nature. Humans are exposed to its spores every day without disease becoming apparent. Only in few cases has *A. niger* been able to colonise the human body as an opportunistic invader and in almost all these cases the patients have a history of severe illness or immunosuppressive treatment.

Animal studies

Several experimental studies to demonstrate the pathogenic potential of *A. niger* have come to the conclusion that neither ingestion of large doses of spores (Nyiredy et al. 1975) nor inhalation of spores (Bhatia and Mohapatra 1969) induces mycosis in experimental animals. One day after ingestion, *A. niger* was no longer detected in the digestive tract, although ingested *A. nidulans* was isolated from the intestine of the animals. In contrast to *A. fumigatus*, which is known to be pathogenic, *A. niger* showed no significant effect on the animals in the inhalation study.

Compromising the immune system by steroid hormones seems to promote the spreading of the fungus in the body after an infection. Jacob et al. (1984) conducted a study with mice infected intravenously with high doses of *A. niger* isolated from sputum. They found evidence of pathogenic action only in groups that had been treated with a hydrocortisone drug. Addition of Decadron, a steroid hormone, to the culture medium of *A. niger* induced more vigorous corneal ulceration in rabbit eyes infected with spores compared to animals inoculated with spores from medium without the steroid (Hasany et al. 1973). The authors conclude that exposure of the ordinarily harmless fungus to the steroid made them behave like a pathogen.

Medical case reports

Washburn et al. (1986) found in their immunological study using human sera that *A. fumigatus* produces substances inhibiting complement, which induces phagocy-

tosis of fungal cells by leukocytes. This defence mechanism against infection was not impaired by liquid from an *A. niger* culture.

Few cases of primary cutaneous aspergillosis caused by *A. niger* are given in the literature. Cahill et al. (1967) reported a severe infection, which had long been wrongly been diagnosed as leprosy, and its successful treatment with nystatin.

Mycosis of the ear is one of the frequent health problems in the tropics. *A. niger* has been isolated from 5% of cases of chronic otitis media in Nigeria (Ibekwe and Okafor 1983). However, the authors consider the fungus to be a secondary invader rather than the causative organism because in most cases the patients had been treated with antibiotics before *A. niger* was isolated from their ears. Paldrok (1965) identified 61 fungal isolates from ear lesions in Sweden, 22 of which were of the *A. niger* group. The high incidence of *A. niger* in the outer ear, they speculate, could possibly be due to the fact that *Aspergilli* are resistant to the fungistatic action of ear wax (cerumen).

These ear infections, called otomycosis, cause local inflammation and mycelial growth on cerumen on the skin of the external ear canal. Whilst relatively common in the tropics, this is not a serious condition and can be treated easily with topical antifungal ointment (Mugliston and O'Donoghue 1985; Paulose et al. 1989). Loh et al. (1998) point to self-cleaning of the ears, leading to mechanical damage of the skin barrier, as an important factor in the occurrence of otomycosis.

A. niger can cause pulmonary infection (Binder et al. 1982; Denning 1998). In rare cases it will invade existing pulmonary cavities and create a ball of matted hyphae known as aspergilloma. This aspergilloma may be present for years and may produce oxalic acid in situ, which may lead to renal problems caused by oxalosis (Nime and Hutchins 1973; Severo et al. 1997).

A number of reports on secondary aspergillosis (reviewed in Abramson et al. 1986; Rippon 1982; Saravia-Gomez 1979), caused by often unidentified *Aspergilli*, describe infection in patients suffering from diabetes, drug abuse, alcoholism, severe diseases such as pneumonia, tuberculosis, enterocolitis or patients receiving antibiotic, steroid, cytotoxin or radiation therapy. These groups of debilitated patients, whose immune systems are in many cases weakened, are characterised by an increased susceptibility to opportunistic micro-organisms that do not pose any risk to healthy people.

While abnormal ports of entry like wounds, burns and lesions of the mucosa can facilitate infection in patients, the only entrances in the healthy person are the digestive and respiratory tracts. Thus, reports on health problems of people extensively exposed to *Aspergillus* spores are of special relevance for risk assessment.

Hypersensitivity

Tomsikova et al. (1981) investigated the hypersensitivity pneumonitis of workers in a citric acid plant. They were exposed to spore dust from the production organism *A. niger* and from contaminating fungi. Although both *Aspergillus* and *Penicillium* have been isolated from the respiratory tracts of this group, the concentration of antibodies against *Penicillium* was significantly higher and more frequent than that against *Aspergillus*. This led the authors to the conclusion that hypersensitivity pneumonitis has developed mainly as a result of inhaled *Penicillium* spores and not *Aspergillus* spores.

In another citric acid plant, Topping et al. (1985) showed that only one-half of the workers suffering from bronchospasm were sensitive to *A. niger* spores, while the other half was sensitive neither to spores nor to other substances collected by filtration from the air inside the plant. Now that the spores have been recognised as the most frequent source of hypersensitivity, their dispersal is minimised by technical means and by turning from surface culture methods to submerged production processes which reduce sporulation of *A. niger*. In an 8-year follow-up study, Seaton and Wales (1994) conclude that *A. niger* is a weak antigen and that simple hygiene measures effectively protect the workforce. If such measures are taken, exclusion of recruits with positive skin tests is then not necessary.

Toxins

Despite the long history and intensive nature of *Aspergillus* research, only few cases of toxin formation by *A. niger* have been reported. However, in no case has *A. niger* been proven to produce aflatoxins or trichotecenes.

Two reports (Kulik and Holaday 1966; Hanssen 1969) that *A. niger* cultures produced aflatoxin B1 have been disproved. The evidence was mainly based on an assay by thin layer chromatography. Chances are that fluorescing substances with similar mobilities (Murakami et al. 1967) have been erroneously interpreted as aflatoxin B1. Later on, more detailed studies (Parrish et al. 1966), including those strains which Kulik and Holaday classified as positive (Mislivec et al. 1968; Wilson et al. 1968), clearly showed that none of the *A. niger* strains produced any aflatoxin. Bullerman and Ayres (1968) were also unable to demonstrate aflatoxin production in *A. niger* they had isolated from cured meats. From the numerous investigations it becomes very clear that *A. niger* does not have the ability to produce aflatoxins.

Several more incidental findings indicate that metabolic products may be toxic (Moreau 1979; Cole and Cox 1981; Reiss 1981). Reports on poisoning of animals after they were fed mouldy feed (Moreau 1979, page 178) are difficult to interpret because toxin formation took place under uncontrolled conditions in storage where various contaminating organisms grow as a mixture. The author suspects oxalic acid, a metabolite of *A. niger*, to be the compound responsible for the toxic effect. Jahn (1977) presents results showing coincidence of the toxic effect of fodder and the presence of an *A. niger* strain producing unusually high amounts of oxalic acid.

The nephrotoxic and carcinogenic mycotoxin ochratoxin A was first reported for the black *Aspergillus* species by Ueno et al. (1991) in the species *A. foetidus*. This was later confirmed by Téren et al. (1996) for another isolate of *A. foetidus*. Abarca et al. (1994) first reported ochratoxin A production in *A. niger* (var. *niger*) and this was later confirmed by Téren et al. (1996), Nakajima et al. (1997), Téren et al. (1997), Heenan et al. (1998) and Taniwaki et al. (1999). Téren et al. (1996) also reported that *A. awamori* produced ochratoxin A, and 1 year previously Ono et al. (1995) reported ochratoxin A production in *A. awamori* var. *fumeus*, *A. amamori* var. *minimus*, *A. usamii* and *A. usamii* mut. shiro-usamii. As all these names are synonyms of *A. niger*, it seems to be confirmed by several authors that some isolates of *A. niger* produces ochratoxin A. However, as mentioned by Varga et al. (2000) only about 6% (1.7% to 18.5% as listed by Abarca et al. 2001) of *A. niger* isolates appear to produce ochratoxin A. More research is needed in order to find out which conditions are optimal for ochratoxin A production by *A. niger* (Frisvad and Samson 2000).

Strain CBS 618.78, listed as *A. foetidus*, when studied under optimal laboratory conditions, has the potential to produce ochratoxin A. CBS 618.78 was regarded as an *A. niger* by Kusters-van Someren et al. (1991). From comparing the several isolates in the culture collections it appears that CBS 618.78 is related to strain CBS 126.48, listed by CBS as *A. niger*, available also as ATCC 10254, NRRL 337, IMI 015954, DSM 734, and IFO 6428. When checked, both strains *A. niger* CBS 126.48 and *A. foetidus* IMI 041871 – identical to *A. foetidus* CBS 618.78 – also produce large amounts of ochratoxin A and B when studied under optimal mycotoxin-inducing conditions. On the other hand, *A. niger* IMI 015954 produced rather small amounts of ochratoxin A. Strain NRRL 337 has been used extensively for production of enzymes (Le Mense et al. 1947; Elmayergi and Scharer 1973; Iwai et al. 1983; Okomura et al. 1983) and organic acids (Shu and Johnson 1947; Bercovitz et al. 1990), and for treatment of baked bean processing wastewater (Hang and Woodams 1979), alcoholic fermentation from alkaline potato peel waste (Bloch et al. 1973), utilisation of brewery spent grain (Hang et al. 1975, 1977) and fungal treatment of beet waste (Hang 1976). In the latter paper the fungus (NRRL 337) is called a "food fungus" and named *A. niger*.

This should be taken as a warning that all *A. niger* van Tieghem isolates, either from nature or obtained from a culture collection should be carefully checked for their potential to produce ochratoxin A at the start of the development of a production process for an enzyme used in the food industry. If a strain exhibits the potential to produce this compound, a control system should be in place to assure that it does not end up in the product at levels that induce a toxic effect. When above-threshold levels are found in a product, the use of the strain in the process should be discontinued (Pariza and Johnson 2001).

A. carbonarius is a more efficient producer of ochratoxin A and a much higher percentage of the isolates tested have been found positive (Horie 1995; Téren et al. 1996; Wicklow et al. 1996; Heenan et al. 1998; Varga et al. 2000). No other species in *Aspergillus* section *Nigri* has been reported to produce ochratoxin A (Varga et al. 2000).

A summary of the ochratoxin A production potential of the species in the *Aspergillus* subgenus *Circumdati* section *Nigri* is given in Table 1.

Kojic acid, though mentioned by Wilson (1971) to be a metabolite of *A. awamori*, is not produced by the *A. niger* strains, as Parrish et al. (1966) clearly stated. This is confirmed by the industrial experience that, under the conditions of enzyme production using *A. niger*, kojic acid has not been demonstrated in the culture liquid.

In a comprehensive screening for toxins from *Aspergilli*, Semeniuk et al. (1971) did not detect any markedly toxigenic strain in the *A. niger* group after they had cultivated 392 *Aspergillus* strains on wheat and soybean feed and fed it to chickens or mice. The authors classified 15 strains of the *A. niger* group (of 34 in the test) as moderately to mildly toxigenic. In *A. awamori*, *A. ellipticus*, *A. heteromorphus* and *A. pulverulentus* they did not find any toxigenicity at all. Unfortunately, the study did not include known toxic substances as controls, which would have been helpful in judging the significance of the results.

In a few cases only, suspected toxins have been purified from cultures and tested in animal studies. Nigragillin has been isolated from cultures of *A. niger* and *A. phoenicis* (Caesar et al. 1969; Cole and Cox 1981, p. 798). The LD_{50} was found to be approximately 250 mg/kg bodyweight when fed to 1-day-old cockerels. Malformins, a group of closely related cyclic peptides, generate deformations in plants and have been isolated from cultures of *A. niger*, *A. ficuum*, *A. awamori* and *A. phoenicis* (Steyn 1977). Two malformins were checked for toxic action on rodents: malformin A1 showed an LD_{50} of 3.1 mg/kg when applied intraperitoneally but there was no evidence of acute toxicity when up to 50 mg/kg were given orally to male mice (Yoshizawa et al. 1975). Anderegg et al. (1976) found the LD_{50} for malformin C in both newborn and 28-day-old rats to be 0.9 mg/kg when it was given intraperitoneally.

A series of naphto-γ-pyrones, produced by some strains of *A. niger*, have been reported to be vertebrate central nervous system toxins (Ghosal et al. 1979; DeLucca et al. 1983; Ehrlich et al. 1984). However, these secondary metabolites cannot be regarded as mycotoxins (Bennett 1989), as they were not shown to be toxic when administered by a natural route but rather after intraperitoneal injection. Furthermore, they accumulate only in the mycelium (Ehrlich et al. 1984). Thus these naphto-γ-pyrones do not appear to be a cause for concern in biotechnological products.

The whole body of knowledge from the literature is carefully taken into consideration when testing industrial strains for any possible risk during the development of a fermentation process. Whenever possible the production

organism is chosen from strains that have been in use for many years and that are examined for their ability to produce known toxins under the fermentation conditions used. Finally, the products are regularly checked to ensure that they meet the requirements of the health authorities as given in the Food Chemical Codex (1996) or in the FAO specifications (FAO/WHO 1992).

Summary

The *A. niger* group is composed of black-spored *Aspergillus* species, several of which have a long history of safe use in the fermentation industry. These species have never been identified to be the primary cause of any disease in man. The risk of allergic hypersensitivity to inhaled spores can be handled in an industrial environment by minimising the exposure of the workers to spore dust. Sporadic toxin formation under undefined conditions has not been observed under controlled fermentation conditions. Thus it is concluded that *A. niger* is a safe production organism for industrial use provided the rules of good manufacturing practice are observed. The relatively new discovery that a low percentage of *A. niger* strains have the potential to produce ochratoxin A under optimal laboratory conditions requires, however, that all *A. niger* isolates, (over)expressing a specific gene of interest, be evaluated for their potential to produce ochratoxin prior to being further developed into a new accepted production strain. The use of strains of established and proven safe industrial strain lineages as hosts to over-express these genes of interest is a good and fast alternative to avoid this.

References

Abarca ML, Bragulat MR, Castellá G, Cabanes FJ (1994) Ochratoxin A production by strains of *Aspergillus niger* var. *niger*. Appl Environ Microbiol 60:2650–2652

Abarca ML, Accensi F, Bragulat MR, Cabanes FJ (2001) Current importance of ochratoxin A producing *Aspergillus* spp. J Food Prot 64:903–906

Abramson C, McCarthy DJ, Reed TM (1986) Morphologic and immunologic implications of pedal cutaneous aspergillosis. J Am Podiatr Med Assoc 76:168–178

Accensi F, Cano J, Figuera L, Abarca ML, Cabanes FJ (1999) New PCR method to differentiate species in the *Aspergillus niger* aggregate. FEMS Microbiol Lett 180:191–196

Al-Mussalam A (1980) Revision of the black *Aspergillus* species. PhD Thesis, Rijksuniversiteit, Utrecht

Anderegg RJ, Biemann K, Büchi G, Cushman M (1976) Malformin C, a new metabolite of *Aspergillus niger*. J Am Chem Soc 98:3365–3370

Ballance DJ, Buxton FP, Turner G (1983) Transformation of *Aspergillus nidulans* by the orotidine-5′-phosphate decarboxylase gene of *Neurospora crassa*. Biochem Biophys Res Commun 112:284–289

Barbesgaard P, Heldt-Hansen HP, Diderichsen B (1992) On the safety of *Aspergillus oryzae*: a review. Appl Microbiol Biotechnol 36:569–572

Bennett JW (1989) Mycotoxins, mycotoxincosis, mycotoxicology and mycopathologica. Mycopathologia 100:3–5

Bercovitz A, Peleg Y, Battat E, Rokem JS, Goldberg I (1990) Localization of pyruvate carboxylase in organic acid-producing *Aspergillus* strains. Appl Environ Microbiol 56:1594–1597

Berka RM, Dunn-Coleman NS, Ward M (1992) Industrial enzymes from *Aspergilli*. In: Bennett JW, Klich MA (eds) *Aspergillus*, biology and industrial applications. Butterworth-Heinemann, London, pp 155–202

Berka RM, Fowler T, Rey MW (1994a) Gene sequence encoding *Aspergillus niger* catalase-R. United States Patent 5,360,901

Berka RM, Fowler T, Rey MW (1994b) Production of *Aspergillus niger* catalase-R. United States Patent 5,360,732

Berufsgenossenschaft der Chemischen Industrie (1998) Merkblatt B 007. Jedermann, Heidelberg

Bhatia YN, Mohapatra LN (1969) Experimental aspergillosis in mice. Part I: Pathogenic potential of *Aspergillus fumigatus*, *Aspergillus flavus and Aspergillus niger*. Mykosen 12:651–654

Binder RE, Faling FJ, Pugatch RD, Mahasaen C, Snider GL (1982) Chronic necrotizing pulmonary aspergillosis: a discrete clinical entity. Medicine 61:109–124

Bloch F, Brown GE, Farkas DF (1973) Utilization of alkaline potato peel waste by fermentation. Amylase production by *Aspergillus foetidus* NRRL 337, and alcoholic fermentation. Am Potato J 50:357–364

Bullerman LB, Ayres JC (1968) Aflatoxin-producing potential of fungi isolated from cured and aged meats. Appl Microbiol 16:1945–1946

Bussink HJD, Kester HCM, Visser J (1990) Molecular cloning, nucleotide sequence and expression of the gene encoding prepro-polygalacturonase II of *Aspergillus niger*. FEBS Lett 273:127–130

Bussink HJD, Buxton FP, Visser J (1991) Expression and sequence comparison of the *Aspergillus niger* and *A. tubingensis* genes encoding polygalacturonase II. Curr Genet 19:467–474

Buxton FP, Gwynne DI, Davies RW (1985) Transformation of *Aspergillus niger* using the *arg*B gene of *Aspergillus nidulans*. Gene 37:207–214

Caesar F, Jansson K, Mutschler E (1969) Nigragilline, a new alkaloid from the *Aspergillus niger* group. I. Isolation and structural elucidation of nigragilline and dioxopiperazine. Helv Physiol Pharmacol Acta 44:676–690

Cahill KM, El Mofty AM, Kawaguchi TP (1967) Primary cutaneous aspergillosis. Arch Dermatol 96:545–547

Campbell EI, Unkles SE, Macro JA, Van den Hondel C, Contreras R, Kinghorn JR (1989) Improved transformation efficiency of *Aspergillus niger* based on the *pyr*G gene. Mol Gen Genet 206:71–75

Cole RJ, Cox RH (1981) Handbook of toxic fungal metabolites. Academic Press, New York

Cullen D, Gray GL, Wilson LJ, Hayenga KJ, Lamsa MH, Rey MW, Norton S, Berka RM (1987) Controlled expression and secretion of bovine chymosin in *Aspergillus nidulans*. Biotechnology 5:369–376

Das A, Roy P (1978) Improved production of citric acid by a diploid strain of *Aspergillus niger*. Can J Microbiol 24:622–625

DeLucca AJ, Ehrlich KB, Ciegler A (1983) Toxicity of extracts of *Aspergillus niger* isolated from stored cotton seed. J Food Safety 5:95–101

Denning DW (1998) Invasive aspergillosis. Clin Infect Dis 26:781–805

Dunn-Coleman NS, Bloebaum P, Berka RM, Bodie E, Robinson N, Armstrong G, Ward M, Przetak M, Carter G, LaCost R, Wilson LJ, Kodama KH, Baliu EF, Bower B, Lamsa M, Heinsohn H (1991) Commercial levels of chymosin production by *Aspergillus*. Biotechnology 9:976–981

Ehrlich KB, DeLucca AJ, Ciegler A (1984) Naphtho-γ-pyrone production by *Aspergillus niger* isolated from stored cottonseed. Appl Environ Microbiol 48:1–4

Elmayergi H, Scharer JM (1973) Physiological studies on *Aspergillus niger* fermentation with polymer additive. J Gen Appl Microbiol 19:385–392

FAO/WHO (1972) Specifications for the identity and purity of some enzymes and certain other substances. 15th report of the Joint FAO/WHO Expert Committee on Food Additives, Food and Agricultural Organization of the United Nations, World Health Organization, Rome

FAO/WHO (1978) Specifications for identity and purity. The 22nd meeting of the Joint FAO/WHO Expert Committee on Food Additives, Food and Agricultural Organization of the United Nations, World Health Organization, Rome

FAO/WHO (1981) Specifications for identity and purity. The 25th meeting of the Joint FAO/WHO Expert Committee on Food Additives, Food and Agricultural Organization of the United Nations, World Health Organization, Rome

FAO/WHO (1987) Toxicological evaluation of certain food additives. The 31st meeting of the Joint FAO/WHO Expert Committee on Food Additives, World Health Organization, Geneva

FAO/WHO (1990) Toxicological evaluation of certain food additives. The 35th meeting of the Joint FAO/WHO Expert Committee on Food Additives, World Health Organization, Geneva

FAO/WHO (1992) Compendium of food additive specifications, vol 1 and 2. Food and Nutrition Paper no. 51/1 and 51/2. Food and Agricultural Organization of the United Nations, World Health Organization, Rome

Federal Register (1993) Direct food substances affirmed as generally recognized as safe; chymosin enzyme preparation derived from *Aspergillus niger* van Tieghem variety *awamori* (Nakazawa) Al-Mussalam. 58:27197

Food Chemical Codex (1996) 4th edn. National Academy Press, Washington

Frisvad JC, Samson RA (2000) *Neopetromyces* gen. nov. and an overview of teleomorphs of *Aspergillus* subgenus *Circumdati*. Stud Mycol (Baarn) 45:201–207

Frisvad JC, Hawksworth DL, Kozakiewicz Z, Pitt JI, Samson RA, Stolk AS (1990) Proposals to conserve important species names in *Aspergillus* and *Penicillium*. In: Samson RA, Pitt JI (eds) Modern concepts in *Penicillium* and *Aspergillus classification*. Plenum Press, New York, pp 83–88

Frost GM, Moss DA (1987) Production of enzymes by fermentation. In: Rehm HJ, Reed G (eds) Biotechnology, vol 7a. VCH, Weinheim, pp 65–102

Gams W, Christensen M, Onions AH, Pitt JI, Samson RA (1985) Intrageneric taxa of *Aspergillus*. In: Samson RA, Pitt JI (eds) Advances in *Penicillium* and *Aspergillus* systematics. Plenum Press, pp 55–62

Ghosal S, Biswas K, Chakrabarti DK (1979) Toxic naphtho-γ-pyrones from *Aspergillus niger*. J Agric Food Chem 27:1347–1351

Gielkens MMC, Visser J, de Graaf LH (1997) Arabinoxylan degradation by fungi: characterization of the arabinoxylan-arabinofuranohydrolase encoding genes of *Aspergillus niger* and *A. tubingensis*. Curr Genet 31:22–29

Godfrey T, Reichelt J (1983) Industrial enzymology. The application of enzymes in industry. Nature Press, New York

Graaf LH de, van den Broek HC, van Ooijen AJJ, Visser J (1994) Regulation of the xylanase-encoding *xlnA* gene of *Aspergillus tubingensis*. Mol Microbiol 12:479–490

Grassin C, Fauguenbergue P (1999) Enzymes, fruit juice processing. In: Flickinger MC, Drew SW (eds) Encyclopedia of bioprocess technology: fermentation, biocatalysis, and bioseparation. Wiley & Sons, Inc, New York

Groot GSP, Herweijer MA, Simonetti ALM, Selten GCM, Misset O (2000) Enzymes in food and feed: past, present and future. In: Bielecki S, Tramper J, Polak J (eds) Food Biotechnology – progress in biotechnology, vol. 17. Elsevier, Amsterdam, pp 95–99

Hang YD (1976) Fungal treatment of beet waste. Prog Water Technol 8:325–327

Hang YD, Woodams EE (1979) Characterization of baked bean processing wastewater and its assimilation by *Aspergillus foetidus*. J Food Sci 44:1548–1549

Hang YD, Splittstoesser DF, Woodams EE (1975) Utilization of brewery spent grain liquor by *Aspergillus niger*. Appl Microbiol 30:879–880

Hang YD, Splittstoesser DF, Woodams EE, Sherman RM (1977) Citric acid fermentation of brewery waste. J Food Sci 42:383–384

Hanssen E (1969) Schädigung von Lebensmitteln durch Aflatoxin B1. Naturwissenschaften 56:90

Harmsen JAM, Kusters-van Someren MA, Visser J (1990) Cloning and expression of a second *Aspergillus niger* pectin lyase gene (pelA): indications of a pectin-lyase gene family in *A. niger*. Curr Genet 18:161–166

Hasany SM, Basu MB, Kazdan MD (1973) Production of corneal ulcer by opportunistic and saprophytic fungi. 1. The effect of pre-treatment of fungi with steroid. Can J Ophthalmol 8:119–131

Heenan CN, Shaw KJ, Pitt JI (1998) Ochratoxin A production by *Aspergillus carbonarius* and *A. niger* isolates and detection using coconut cream agar. J Food Mycol 1:67–72

Horie Y (1995) Productivity of ochratoxin of *Aspergillus carbonarius* in *Aspergillus* section *Nigri*. Nippon Kingakkai Kaiho 36:73–76

Ibekwe AC, Okafor JI (1983) Pathogenic organisms in chronic suppurative otitis media in Enugu, Nigeria. Trop Geogr Med 35:389–391

Iwai M, Okumura S, de Leal EL, Tsujisaka Y (1983) Purification of four esterases from *Aspergillus niger* NRRL 337. Agric Biol Chem 47:1865–1868

Jacob Z, Shukla PK, Wahab S, Srivastava OP, Srivastava GN (1984) *Aspergillus terreus, A. fumigatus* and *A. niger* from the sputum of a human patient, their growth, drug sensitivity and pathogenicity. Biol Mem 9:87–94

Jahn F (1977) Oxalsäurebildungsvermögen und Toxizität eines Stammes von *Aspergillus niger* aus Verschimmeltem Heu. Wien Tierärztl Monatsschr 64:79–82

Kelly JM, Hynes MJ (1985) Transformation of *Aspergillus niger* by the amdS gene of *Aspergillus nidulans*. EMBO J 4:475–479

Kluyver AJ, Perquin LHC (1932) Zur Methodik der Schimmelstoffwechseluntersuchung. Biochem Z 266:68–81

Kozakiewicz Z (1989) *Aspergillus* species on stored products. Mycol Pap 161:1–188

Kozakiewicz Z, Frisvad JC, Hawksworth DL, Pitt JI, Samson RA, Stolk AC (1992) Proposals for nomina specifica conservanda and rejicienda in *Aspergillus* and *Penicillium* (Fungi). Taxon 41:109–113

Kulik MM, Holaday CE (1966) Aflatoxin: a metabolic product of several fungi. Mycopathol Mycol Appl 30:137–140

Kusters-van Someren MA, Kester HMC, Samson RA, Visser J (1990) In: Samson RA, Pitt JI (eds) Modern concepts in *Penicillium* and *Aspergillus* classification. Plenum Press, New York, pp 83–88

Kusters-van Someren MA, Samson RA, Visser J (1991) The use of RFLP analysis in classification of the black *Aspergilli*: reinterpretation of the *Aspergillus niger* aggregate. Curr Genet 19:21–26

Le Mense EH, Corman J, van Lanen JM, Langlykke AF (1947) Production of mold amylases in submerged culture. J Bacteriol 54:149–159

Loh KS, Tan KK, Kumarasinghe G, Leong HK, Yeoh KH (1998) Otitus externa – the clinical pattern in a tertiary institution in Singapore. Ann Acad Med Singapore 27:215–218

Megnegneau B, Debets F, Hoekstra RF (1993) Genetic variability and relatedness in the complex group of black Aspergilli based on random amplification of polymorphic DNA. Curr Genet 23:323–329

Mislivec PB, Hunter JH, Tuite J (1968) Assay for aflatoxin production by the genera *Aspergillus* and *Penicillium*. Appl Microbiol 16:1053–1055

Moreau C (1979) Moulds, toxins and food. Wiley, New York

Mugliston T, O'Donoghue G (1985) Otomycosis – a continuing problem. J Laryngol Otol 99:327–333

Mullaney EJ, Gibson DM, Ullah AHJ (1991) Positive identification of a lambda gtll clone containing a region of fungal phytase gene by immunoprobe and sequence verification. Appl Microbiol Biotechnol 35:611–614

Murakami H, Takase S, Ishii T (1967) Non-productivity of aflatoxin by Japanese industrial strains of *Aspergillus*. I. Production of fluorescent substances in agar slant and shaking cultures. J Gen Appl Microbiol 13:323–334

Nakajima M, Tsubouchi H, Miyabe M, Ueno Y (1997) Survey of aflatoxin B1 and ochratoxin A in commercial green coffee beans by high-performance liquid chromatography linked with immunoaffinity chromatography. Food Agric Immunol 9:77–83

Nguyen QK, Ruttkowski E, Leidinger K, Albrecht H, Gottschalk M (1991) Characterization and expression of a genomic pectin methyl esterase-encoding gene in *Aspergillus niger*. Gene 106:71–77

Nime FA, Hutchins GM (1973) Oxalosis caused by *Aspergillus* infection. Johns Hopkins Med J 133:183–194

Nunberg J, Schweikart V, Lawyer F, Innis M, Flatgaar J, Wittman W, Meade J (1983) Cloning of the *Aspergillus awamori* glucoamylase gene. Proceedings of the ASM conference on gene manipulations in the exploitation and study of fungi. Elkhart, Indiana

Nunberg JH, Meade JH, Cole G, Lawyer FC, McCabe P, Schweickart V, Tal R, Wittman VP, Flatgaard JE, Innes MA (1984) Molecular cloning and characterization of the glucoamylase gene of *Aspergillus awamori*. Mol Cell Biol 4:2306–2315

Nyiredy I, Etter L, Fesüs I, Mayer G (1975) The fate of mould "spores" in the digestive tract of chicks. Acta Vet Acad Sci Hung 25:123–128

Okumura S, Iwai M, Tsujisaka Y (1983) Properties and substrate specificities of four esterases from *Aspergillus niger* NRRL 337. Agric Biol Chem 47:1869–1872

Ono H, Kaqtaoka A, Koakutsu M, Tanaka K, Kawasugi S, Wakazawa M, Ueno Y, Manabe M (1995) Ochratoxin A producibility by strains of *Aspergillus niger* group stored in IFO culture collection. Mycotoxins 41:47–51

Paldrok H (1965) Report on a case of subcutaneous dissemination of *Aspergillus niger*, type *awamori*. Acta Derm Venereol 45:275–282

Parenicová L, Suykerbuyk MEG, Samson RA, Visser J (1997) Evaluation of RFLP analysis for the classification of selected black Aspergilli. Mycol Res 101:810–814

Parenicová L, Skouboe P, Samson RA, Rossen L, Visser J (2000) Molecular tools for the classification of black aspergilli. In: Samson RA, Pitt JI (eds) Integration of modern taxonomic methods for *Penicillium* and *Aspergillus* classification. Harwood, Amsterdam, pp 413–424

Parenicová L, Skouboe P, Frisvad JC, Samson RA, Rossen L, ten Hoor-Suykerbuyk M, Visser J (2001) Combined molecular and biochemical approach identifies *Aspergillus japonicus* and *Aspergillus aculeatus* as two species. Appl Environ Microbiol 67:521–527

Pariza MW, Johnson EA (2001) Evaluating the safety of microbial enzyme preparations used in food processing: update for a new century. Regul Toxicol Pharmacol 33:1–14

Parrish FW, Wile BJ, Simmons EG, Long L Jr (1966) Production of aflatoxins and kojic acid by species of *Aspergillus* and *Penicillium*. Appl Microbiol 14:139

Paulose KO, Alkhalifa S, Shenoy P, Sharma RK (1989) Mycotic infection of the ear (otomycosis): a prospective study. J Laryngol Otol 103:30–35

Piddington CS, Houston CS, Paloheimo M, Cantrell M, Miettinen-Oinonen A, Nevalainen H, Rambosek J (1993) The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5 optimum acid phosphatase (aph) from *Aspergillus niger* var. *awamori*. Gene 133:55–62

Pitt JI, Samson RA (2000) Types of *Aspergillus* and *Penicillium* and their teleomorphs in current use. In: Samson RA, Pitt JI (eds) Integration of modern taxonomic methods for *Penicillium* and *Aspergillus* classification. Harwood, Amsterdam, pp 51–72

Pitt JI, Swamson RA, Frisvad JC (2000) List of accepted species and their synonyms in the family Trichocomaceae. In: Samson RA, Pitt JI (eds) Integration of modern taxonomic methods for *Penicillium* and *Aspergillus* classification. Harwood, Amsterdam, pp 9–49

Raper KB, Fennell DI (1965) The genus *Aspergillus*. Williams and Wilkins, Baltimore

Reiss J (1981) Mycotoxine in Lebensmitteln. Fischer, Stuttgart

Reiss J (1986) Schimmelpilze. Lebensweise, Nutzen, Schaden, Bekämpfung. Springer, Berlin Heidelberg New York, pp 33–41

Rippel-Baldes A (1955) Grundzüge der Mikrobiologie, 3rd edn. Springer, Berlin Heidelberg New York

Rippon JW (1982) Medical mycology: the pathogenic fungi and the pathogenic actinomycetes, 2nd edn. Saunders, Philadelphia, pp 565–594

Roukas T (2000) Citric and gluconic acid production from fig by *Aspergillus niger* using solid state fermentation. J Ind Microbiol Biotechnol 25:298–304

Samson RA (1992) Current taxonomic schemes of the genus *Aspergillus* and its teleomorphs. In: Bennett JW, Klich MA (eds) *Aspergillus*, biology and industrial applications. Butterworth-Heinemann, London, pp 355–390

Saravia-Gomez J (1979) Aspergillosis of the central nervous system. In: Vinken PJ, Bruyen GW (eds) Handbook of clinical neurology, vol 35. Infections of nervous systems, part III. GW North Holland, pp 395–400

Seaton A, Wales D (1994) Clinical reactions to *Aspergillus niger* in a biotechnology plant: an eight year follow up. Occup Environ Med 51:54–56

Selten G (1994) The versatile *Aspergillus niger*. Gist 60:5–7

Selten GCM, van Gorcom RFM, Swinkels BW (1995) Selection marker gene free recombinant strains: a method for obtaining them and the use of these strains. European Patent Application EPO0635574 A1

Selten GCM, Swinkels BW, Bovenberg RAL (1998) Gene conversion as a tool for the construction of recombinant industrial filamentous fungi. International Patent Application WO 98/46772

Semeniuk G, Harshfield GS, Carlson CW, Hesseltine CW, Kwolek WF (1971) Mycotoxins in *Aspergillus*. Mycopathol Mycol Appl 43:137–152

Severo LC, Geyer GR, Da Silva Porto N, Wagner MB, Londero AT (1997) Pulmonary *Aspergillus niger* colonisation. Report of 23 cases and a review of the literature. Rev Iberoam Micol 14:104-110

Shu P, Johnson MJ (1947) Effect of the composition of the sporulation medium on citric acid production *by Aspergillus niger* in submerged culture. J Bacteriol 54:161–167

Steyn PS (1977) Some new mycotoxins. Pure Appl Chem 49:1771–1778

Taniwaki MH, Pitt JI, Urbano GR, Teixeira AA, Leitao MFF (1999) Fungi producing ochratoxin A in coffee. 18th International Scientific Colloquium on Coffee. ASIC, Helsinki, pp 239–247

Téren J, Varga J, Hamari Z, Rinyu E, Kevei F (1996) Immunochemical detection of ochratoxin A in black *Aspergillus* strains. Mycopathologia 134:171–176

Téren J, Palágyi A, Varga J (1997) Isolation of ochratoxin producing aspergilli from green coffee beans of different origin. Cereal Res Commun 25:303–304

Tilburn J, Scazzocchio C, Taylor GG, Zabicky-Zissman JH, Lockington RA, Davis RW (1983) Transformation by integration in *Aspergillus nidulans*. Gene 26:205–221

Tomsikova A, Zavazal V, Sach J, Novackova D (1981) Ein der Farmerlunge Ähnliches Syndrom. Mykosen 24:354–366

Topping MD, Scarisbrick DS, Luczynska CM, Clarke EC, Seaton A (1985) Clinical and immunological reactions to *Aspergillus niger* among workers at a biotechnology plant. Brit J Ind Med 42:312–318

Ueno Y, Kawamura O, Sugiura Y, Horiguchi K, Nakajima M, Yamamoto K, Sato S (1991) Use of monoclonal antibodies, enzyme-linked immunosorbent assay and immunoaffinity column chromatography to determine ochratoxin A in porcine sera, coffee products and toxin-producing fungi. In: Castagnero M, Plestina R, Dirheimer G, Chernozemsky IN, Bartsch H (eds) Mycotoxins, endemic nephropathy and urinary tract tumors. IARC Scientific Publication No. 115. International Agency for Research on Cancer, Lyon, pp 71–75

Van Dijck PWM (1999) Chymosin and phytase. Made by genetic engineering. J Biotechnol 67:77–80

Van Gorcom RFM, van Hartingsveldt W, van Paridon PA, Veenstra AE, Luiten RGM, Selten GCM (1991) Cloning and expression of microbial phytase. European Patent Application 0.420.358

Van Hartingsveldt W, Mattern IE, Van Zeijl CMJ, Pouwels PH, van den Hondel CAMJJ (1987) Development of a homologous transformation system for *Aspergillus niger* based on the *pyr*G gene. Mol Gen Genet 206:1–75

Van Hartingsveldt W, van Zeijl CMJ, Harteveld GM, Gouka RJ, Suykerbuyk MEG, Luiten RGM, van Paridon PA, Selten GCM, Veenstra AE, van Gorcom RFM, van den Hondel CAMJJ (1993) Cloning, characterization and overexpression of the phytase-encoding gene (*phy*A) of *Aspergillus niger*. Gene 127:87–94

Varga J, Kevei F, Fekete C, Coenen A, Kozakiewicz Z, Croft JH (1993) Restriction fragment length polymorphisms in the mitochondrial DNAs of the *Aspergillus niger* aggregate. Mycol Res 97:1207–1212

Varga J, Kevei F, Vriesma A, Debets F, Kozakiewicz Z, Croft JH (1994) Mitochondrial DNA restriction fragment length polymorphisms in field isolates of the *Aspergillus niger* aggregate. Can J Microbiol 40:612–621

Varga J, Kevei F, Hamari Z, Tóth B, Téren J, Croft JH, Kozakiewicz Z (2000) Genotypic and phenotypic variability among black aspergilli. In: Samson RA, Pitt JI (eds) Integration of modern taxonomic methods for *Penicillium* and *Aspergillus* classification. Harwood, Amsterdam, pp 397–411

Ward M, Wilson LJ, Carmonia CL, Turner G (1988) The *olic*3 gene of *Aspergillus niger* – isolation, sequence and use as a selective marker for transformation. Curr Genet 14:37–42

Washburn RG, Hammer CH, Bennett JE (1986) Inhibition of complement by culture supernatants of *Aspergillus fumigatus*. J Infect Dis 154:944–951

Wicklow DT, Dowd PF, Alfatafta AA, Gloer JB (1996) Ochratoxin A: an anti-insectan metabolite from the sclerotia of *Aspergillus carbonarius* NRRL 369. Can J Microbiol 42:1100–1103

Wilson BJ (1971) Miscellaneous *Aspergillus* toxins. In: Ciegler A, Kadis S, Ajl SJ (eds) Microbial toxins, vol 6. Academic Press, New York, pp 207–295

Wilson BJ, Campbell TC, Hayes AW, Hanlin RT (1968) Investigation of reported aflatoxin production by fungi outside the *Aspergillus flavus* group. Appl Microbiol 16:819–821

Yoshizawa T, Tsuchiya Y, Morooka N, Sawada Y (1975) Malformin A as a mammalian toxicant from *Aspergillus niger*. Agric Biol Chem 39:1325–1326

This report contains the collective views of an international group of experts and does not necessarily represent the decisions or the stated policy of the World Health Organization or of the Food and Agriculture Organization of the United Nations

Evaluation of certain food additives and contaminants

Thirty-fifth Report of the
Joint FAO/WHO Expert Committee on
Food Additives

World Health Organization
Technical Report Series
789

World Health Organization, Geneva 1990

WHO Library Cataloguing in Publication Data

Joint FAO/WHO Expert Committee on Food Additives
Evaluation of certain food additives and contaminants: thirty-fifth report of the Joint FAO/WHO Expert Committee on Food Additives.

(World Health Organization technical report series; 789)

1. Food additives – analysis 2. Food additives – toxicity 3. Food contamination I. Series

ISBN 92 4 120789 2 (NLM Classification: WA 712)
ISSN 0512-3054

PRINTED IN SWITZERLAND
89/8241 – Schüler SA – 6800

CONTENTS

insignificant relative to naturally occurring levels in the diet, and there is no reason to suppose that they present a hazard.

The Committee also reviewed new toxicological studies on a palm-oil sucroglyceride, including a long-term carcinogenicity study in rats and short-term studies in rats and dogs.

It was concluded that, both for sucrose esters of fatty acids manufactured by a process using dimethylsulfoxide, isobutanol, ethyl methyl ketone, or a combination of these as solvent, and for the palm-oil sucroglyceride, the previously established group ADI of 0–10 mg per kg of body weight for sucrose esters of fatty acids and sucroglycerides would apply.

An addendum to the toxicological monograph was prepared.

The specifications for sucrose esters of fatty acids were revised to include considerations on the use of the above-mentioned solvents.

The existing specifications for sucroglycerides were maintained.

3.1.2 *Enzyme preparations*

Enzyme preparations derived from Aspergillus niger

As a consequence of its review of general specifications for enzyme preparations, the Committee reconsidered the evaluation of enzymes derived from *Aspergillus niger* made at the thirty-first meeting (Annex 1, reference 77). At that meeting, the Committee established a single ADI for several separate enzyme preparations derived from *Aspergillus niger* of 0–1 mg of total organic solids per kg of body weight. The enzyme preparations for which this ADI was established were carbohydrases, amyloglucosidases (EC 3.2.1.3), endo-1,3(4)-β-glucanase (EC 3.2.1.6), hemi-cellulase, pectinases (EC 3.1.1.11; 4.2.2.10; 3.2.1.15), and protease.

In view of the fact that *Aspergillus niger* is a common organism in food, that many strains have had a long history of use as an enzyme source, and that the numerous studies of various preparations from various strains have demonstrated no hazard to human health, the numerical ADI that was earlier established for each of the above-listed enzyme preparations from *Aspergillus niger* was changed to an ADI "not specified".

A toxicological monograph was not prepared.

None of the existing specifications for enzyme preparations derived from *Aspergillus niger* were reviewed.

[1] Production of Single-Stranded Plasmid DNA

By JEFFREY VIEIRA and JOACHIM MESSING

Introduction

In the study of gene structure and function, the techniques of DNA analysis that are efficiently carried out on single-strand (ss) DNA templates, such as DNA sequencing and site-specific *in vitro* mutagenesis, have been of great importance. Because of this, the vectors developed from the ssDNA bacteriophages M13, fd, or f1, which allow the easy isolation of strand-specific templates, have been widely used. While these vectors are very valuable for the production of ssDNA, they have certain negative aspects in comparison to plasmid vectors (e.g., increased instability of some inserts, the minimum size of phage vectors). Work from the laboratory of N. Zinder showed that a plasmid carrying the intergenic region (IG) of f1 could be packaged as ssDNA into a viral particle by a helper phage.[1] This led to the construction of vectors that could combine the advantages of both plasmid and phage vectors.[2] Since that time a number of plasmids carrying the intergenic region of M13 or f1 have been constructed with a variety of features.[3]

A problem that has been encountered in the use of these plasmid/phage chimeric vectors (plage) is the significant reduction in the amount of ssDNA that is produced as compared to phage vectors. Phage vectors can have titers of plaque-forming units (pfu) of 10^{12}/ml and give yields of a few micrograms per milliliter of ssDNA. It might then be expected that cells carrying both a plage and helper phage would give titers of 5×10^{11}/ml for each of the two. However, this is not the case due to interference by the plage with the replication of the phage.[4] This results in a reduction in the phage copy number and, therefore, reduces the phage gene products necessary for production of ssDNA. This interference results in a 10- to 100-fold reduction in the phage titer and a level of ss plasmid DNA particles of about 10^{10} colony forming units (cfu) per milliliter.[1] Phage mutants that show interference resistance have been isolated.[4,5] These mutants can increase the yield of ss plasmid by 10-fold and concurrently

[1] G. P. Dotto, V. Enea, and N. D. Zinder, *Virology* **114,** 463 (1981).
[2] N. D. Zinder and J. D. Boeke, *Gene* **19,** 1 (1982).
[3] D. Mead and B. Kemper, *in* "Vectors: A Survey of Molecular Cloning Vectors and Their Uses." Butterworth, Massachusetts, 1986.
[4] V. Enea and N. D. Zinder, *Virology* **122,** 222 (1982).
[5] A. Levinson, D. Silver, and B. Seed, *J. Mol. Appl. Genet.* **2,** 507 (1984).

increase the level of phage by a similar amount. Whether wild-type (wt) phage or an interference-resistant mutant is used as helper the yield of plasmid ssDNA is usually about equal to that of the phage,[3] and as the plasmid size increases the ratio shifts to favor the phage.[5] In order to increase both the quantitative and qualitative yield of the plasmid ssDNA, a helper phage, M13KO7, has been constructed that preferentially packages plasmid DNA over phage DNA. In this chapter, M13KO7 will be described and its uses discussed.

M13 Biology

Certain aspects of M13 biology and M13 mutants play an important role in the functioning of M13KO7, so a short review of its biology is appropriate.[6,7] M13 is a phage that contains a circular ssDNA molecule of 6407 bases packaged in a filamentous virion which is extruded from the cell without lysis. It can infect only cells having an F pili, to which it binds for entering the cell. The phage genome consists of 9 genes encoding 10 proteins and contains an intergenic region of 508 bases. The proteins expressed by the phage are involved in the following processes: I and IV are involved in phage morphogenesis, III, VI, VII, VIII, and IX are virion proteins, V is an ssDNA binding protein, X is probably involved in replication, and II creates a site-specific (+) strand nick within the IG region of the double-stranded replicative form (RF) of the phage DNA molecule at which DNA synthesis is initiated.

Phage replication consists of three phases: (1) ss–ds, (2) ds–ds, and (3) ds–ss. The ss–ds phase is carried out entirely by host enzymes. For phases 2 and 3, gene II, which encodes both proteins II and X, is required for initiating DNA synthesis; all other functions necessary for synthesis are supplied by the host. The DNA synthesis initiated by the action of the gene II protein (gIIp) leads to both the replication of the ds molecule and the production of the ssDNA that is to be packaged in the mature virion. The phage is replicated by a rolling circle mechanism that is terminated by gIIp cleaving the displaced (+) strand at the same site and resealing it to create a circular ssDNA molecule. Early in the phage life cycle this ssDNA molecule is converted to the ds RF but later in the phage life cycle gVp binds to the (+) strand, preventing it from being converted to dsDNA and resulting in it being packaged into viral particles. The assembly of the virion occurs in the cell membrane where the gVp is replaced by the

[6] D. T. Denhardt, D. Dressler, and D. S. Ray (eds.), "The Single-Stranded DNA Phages." Cold Spring Harbor Lab., Cold Spring Harbor, New York, 1978.
[7] N. D. Zinder and K. Horiuchi, *Microbiol. Rev.* **49**, 101 (1985).

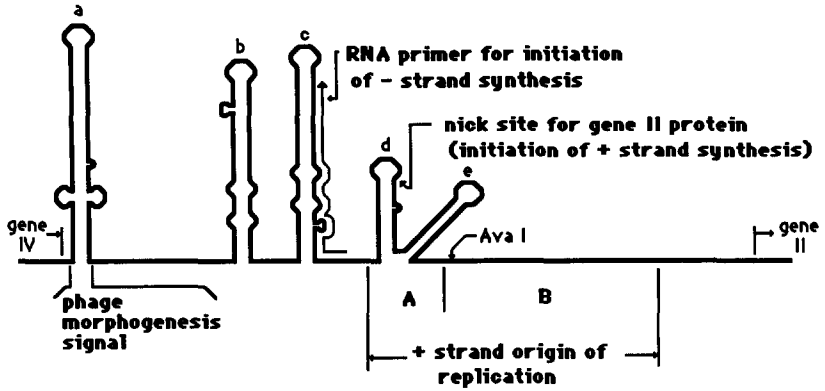
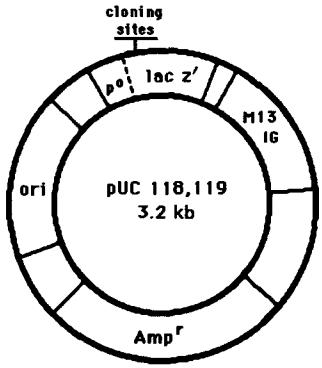


FIG. 1. The M13 intergenic region is schematically presented. It is 508 nucleotides long and is situated between genes II and IV. Potential secondary structure is represented by hairpin structures a–e.[8] Important functional regions are also shown.

gVIIIp and the other virion proteins as the phage particle is extruded from the cell.

The IG structure contains regions important for four phage processes[8–10]: (1) The sequences necessary for the recognition of an ssDNA by phage proteins for its efficient packaging into viral particles; (2) the site of synthesis of an RNA primer that is used to initiate (−) strand synthesis; (3) the initiation; and (4) the termination of (+) strand synthesis. In Fig. 1 the IG, which has the potential to form five hairpin structures, is represented schematically and important regions designated. Most important to the functioning of M13KO7 is the origin of replication of the (+) strand. The origin consists of 140 bp and can be divided into two domains. Domain A, about 40 bp, is essential for replication and contains the recognition sequence for gIIp to create the nick that initiates and terminates replication of the RF. Domain B is about 100 bp long and acts as an enhancer for gIIp to function at domain A. The effect of domain B can be demonstrated by the fact that a disruption or deletion of it will decrease phage yield by 100-fold.[9] Two types of mutants, a qualitative mutation from M13mp1[11] and two quantitative ones from R218 and R325,[12] that compensate for the loss of a functional domain B have been analyzed. The qualitative mutant from mp1, which has an 800-bp insertion within B,

[8] H. Schaller, *Cold Spring Harbor Symp. Quant. Biol.* **45,** 177 (1978).
[9] G. P. Dotto, K. Horiuchi, and N. D. Zinder, *J. Mol. Biol.* **172,** 507 (1984).
[10] G. P. Dotto and N. D. Zinder, *Virology* **130,** 252 (1983).
[11] J. Messing, B. Gronenborn, B. Muller-Hill, and P. H. Hofschneider, *Proc. Natl. Acad. Sci. U.S.A.* **74,** 3642 (1977).
[12] G. P. Dotto and N. D. Zinder, *Proc. Natl. Acad. Sci. U.S.A.* **81,** 1336 (1984).



Cloning Sites

pUC 118

```
                                  Xma I
met-->lac z'          Sst I       Sma I       Xba I       Pst I       HindIII
ATGACCATGATTACGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGCA
            EcoRI       Kpn I     BamHI       Sal I       Sph I
                        Asp718                Acc I
                                              HincII
```

pUC 119

```
                                  HincII
                                  Acc I                 Asp718
met-->lac z'          Sph I       Sal I       BamHI     Kpn I       EcoRI
ATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTC
            HindIII     Pst I       Xba I       Sma I     Sst I
                                                Xma I
```

FIG. 2. Structure of pUC 118 and 119 and the DNA sequence of the unique restriction enzyme sites within the sequence encoding the *lacZ* peptide.

consists of a single G-to-T substitution that changes a methionine (codon 40) to an isoleucine within the gIIp.[13] This change allows the mp1gIIp to function efficiently enough on an origin consisting of only domain A to give wild-type levels of phage. In R218 and R325 the loss of a functional domain B is compensated for by mutations that cause the overproduction of a normal gIIp at 10-fold normal levels.[12,13] Even though a wild-type gIIp works very poorly on a domain B-deficient origin, the excess level of gIIp achieves enough initiation of replication to give normal levels of phage.

pUC 118 and 119

All ss plasmid DNA vectors carry a phage intergenic region. The entire complement of functions necessary for the packaging of ssDNA

[13] G. P. Dotto, K. Horiuchi, and N. D. Zinder, *Nature (London)* **311,** 279 (1984).

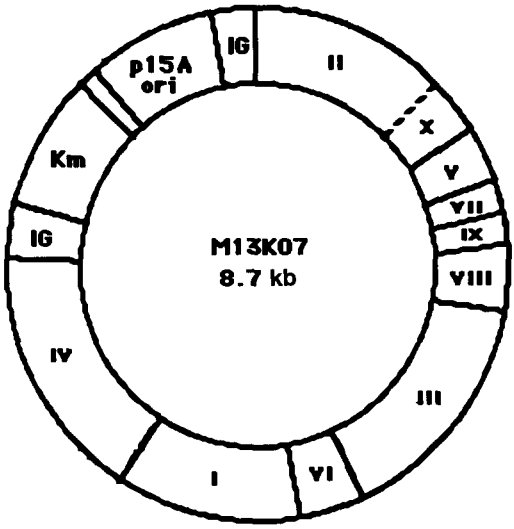


FIG. 3. Structure of M13KO7.

into viral particles will work *in trans* on an IG region. The vectors used in the experiments described here are pUC 118 and 119 (Fig. 2). They are pUC 18 and 19,[14] respectively, with the IG region of M13 from the *Hgi*AI site (5465) to the *Dra*I site (5941) inserted at the unique *Nde*I site (2499) of pUC. The orientation of the M13 IG region is such that the strand of the *lac* region that is packaged as ssDNA is the same as in the M13mp vectors.

M13KO7

M13KO7 (Fig. 3) is an M13 phage that has the gene II of M13mp1 and the insertion of the origin of replication from p15A[15] and the kanamycin-resistance gene from Tn 903[16] at the *Ava*I site (5825) of M13. With the p15A origin, the phage is able to replicate independent of gIIp. This allows the phage to overcome the effects of interference and maintain adequate genome levels for the expression of proteins needed for ssDNA production when it is growing in the presence of a plage. The effect of the addition of the plasmid origin is shown in Fig. 4B. The insertion of the p15A origin and the kanamycin-resistance gene separates the A and B

[14] J. Norrander, T. Kempe, and J. Messing, *Gene* **26,** 101 (1983).
[15] G. Selzer, T. Som, T. Itoh, and J. Tomizawa, *Cell* **32,** 119 (1983).
[16] N. D. F. Grindley and C. M. Joyce, *Proc. Natl. Acad. Sci. U.S.A.* **77,** 7176 (1980).

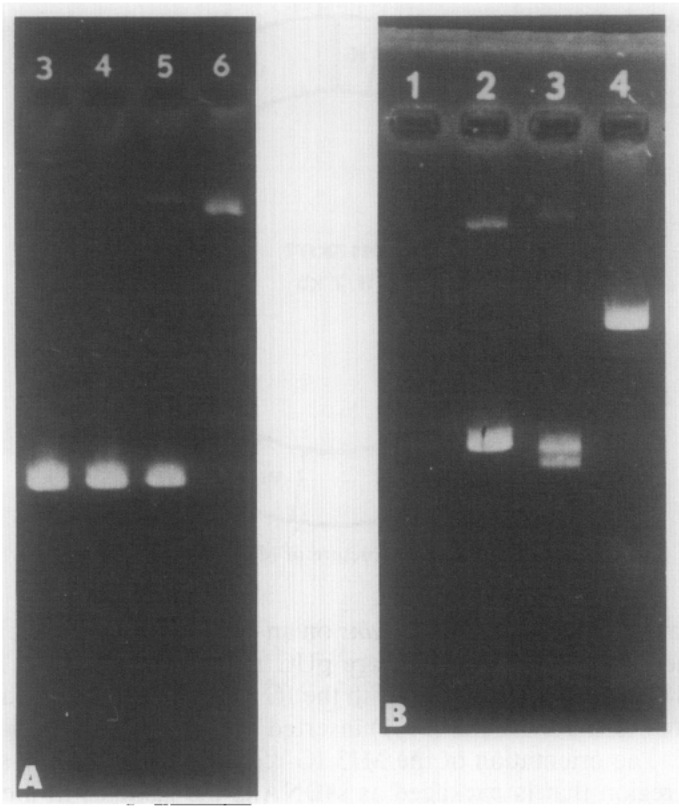


FIG. 4. In all gel lanes 40 μl of the supernatant fraction after centrifugation of the culture was mixed with 6 μl of SDS gel-loading buffer and loaded on the gel. (A) Lane 3: pUC 118 with M13KO7 as helper phage. Plasmid titer is 5×10^{11} cfu/ml, phage titer is 8×10^{9} pfu/ml. Lane 4: pUC 119 with M13KO7 as helper phage. Plasmid titer is 6×10^{11} cfu/ml, phage titer is 8×10^{9} pfu/ml. Lane 5: pUC 119 with M13KO19 (similar to KO7, but with a deletion of domain B of the phage origin of replication) as helper phage. Lane 6: M13KO7. (B) Lane 1: pUC 119 with an M13mp8 phage carrying the kanamycin gene, but no plasmid origin of replication, as helper phage. Lane 2: pUC 119 with M13KO19 as helper phage. Lane 3: pUC 19 with the M13 IG region in the same location as 119, but in the opposite orientation. Lane 4: pUC 118 with 2.5-kb insert.

domains of the phage origin of replication, creating an origin that is less efficient for the functioning of the mp1 gIIp than the wild-type origin carried by the plage. This, plus the high copy number of pUC, leads to the preferential packaging of plasmid DNA into viral particles. The mp1 gIIp functions well enough on the altered origin when M13KO7 is grown by itself to produce a high titer of phage for use as inoculum for the production of ss plasmid.

Materials and Reagents

Strains

MV1184: *ara*,Δ(*lac–pro*), *strA*, thi, (ϕ80Δ*lacIZ*ΔM15),Δ(*srl–recA*) 306::Tn*10*(tet^r); F′: *traD*36, *proAB*, *lacI*^{q}ZΔm15)

Media

2× YT (per liter): 16 g Difco Bacto tryptone, 10 g Difco Bacto yeast extract, 5 g NaCl, 10 m*M* KPO_4, pH 7.5

2× YT plates: 15 g Difco Bacto agar added to 1 liter of 2× YT

YT soft agar (per liter): 8 g Difco Bacto tryptone, 5 g yeast extract, 5 g NaCl, 7 g agar

M9 plates: For 1 liter of 10× M9 salts: combine 60 g Na_2HPO_4, 30 g KH_2PO_4, 0.5 g NaCl, 10 g NH_4Cl dissolved in H_2O to a final volume of 970 ml and autoclave. After autoclaving add 10 ml of a sterile 1 *M* $MgSO_4$ solution and 20 ml of a sterile 0.05 *M* $CaCl_2$ solution. For 1 liter of plates autoclave 15 g of agar in 890 ml. After autoclaving add 100 ml 10× M9 salts, 10 ml of a 20% glucose solution, and 1 ml of a 1% thiamin solution

Solutions

SDS gel loading buffer: 0.05% bromphenol blue, 0.2 *M* EDTA, pH 8.0, 50% glycerol, 1% SDS

TE buffer: 10 m*M* Tris–HCl, pH 8.0, 1 m*M* EDTA, pH 8.0

Growth of M13KO7

M13KO7 exhibits some instability of the insert during growth, but this does not create a problem if it is propagated correctly. The procedure for the production of M13KO7 is the following. M13KO7 supernatant is streaked on a YT agar plate and then 4 ml of soft agar, to which 0.5 ml of a culture of MV 1184 ($OD_{600} > 0.8$) has been added, is poured across the plate from the dilute side of the streak toward the more concentrated side. After 6–12 hr of incubation at 37° single plaques are picked and grown individually in 2–3 ml of YT containing kanamycin (70 μg/ml) overnight. The cells are then pelleted by centrifugation, and the supernatant is used as inoculum of M13KO7. The phage in the supernatant will remain viable for months when stored at 4°.

Production of ss Plasmid DNA

For the production of ss plasmid DNA it is important that a low-density culture of plage-containing cells, infected with M13KO7, be grown for 14–18 hr with very good aeration. The medium that is used is

2× YT supplemented with 0.001% thiamin, 150 μg/ml ampicillin, and, when appropriate, 70 μg/ml kanamycin. Commonly used methods are the following:

1. A culture of MV1184 (pUC 118/119) in early log phase is infected with M13KO7 at a multiplicity of infection (moi) of 2–10 and incubated at 37° for 1 hr and 15 min. The infection should be carried out on a roller or a shaker at low rpm. After this time the cells are diluted, if necessary, to an $OD_{600} < 0.2$ and kanamycin is added to a final concentration of 70 μg/ml. The culture is then grown for 14–18 hr at 37°. Culture conditions are usually 2–3 ml in an 18-mm culture tube on a roller or 5–10 ml in a 125-ml culture flask on a shaker at 300 rpm. Pellet the cells by centrifugation (8000 *g*, 10 min) and remove the supernatant to a fresh tube. Add one-ninth of the supernatant volume of 40% PEG and of 5 *M* sodium acetate and mix well. Place on ice 30 min and pellet the viral particles by centrifugation (8000 *g*, 10 min) and pour off the supernatant. Remove the remaining supernatant with a sterile cotton swab. Resuspend the pellet in 200 μl TE buffer by vortexing. Add 150 μl of TE-saturated phenol (pH 7) and vortex for 30 sec. Add 50 μl of $CHCl_3$, vortex, and centrifuge for 5 min (Brinkman Eppendorf centrifuge). Remove the aqueous layer to a fresh tube and repeat phenol/$CHCl_3$ extraction. Remove the aqueous layer to a fresh tube and add an equal volume of $CHCl_3$, vortex, and centrifuge for 5 min. Remove the aqueous layer to another tube and add 3 vol of ether. Vortex well and centrifuge briefly. Remove the ether, add one-twentieth the volume of 3 *M* sodium acetate (pH 7), and precipitate the DNA with 2.5 vol of ethanol at −70° for 30 min and then pellet by centrifugation. Once the pellet is dry it can be resuspended in TE and used in the same manner as has been previously described for the use of M13 ssDNA templates.[17]

2. For the screening of plasmid for inserts a colony selected from a plate is added to 2–3 ml of medium containing M13KO7 ($\sim 10^7$/ml) and grown at 37° for a few hours. Kanamycin is then added and the cultures are incubated for 14–18 hr at 37°. The cells are then pelleted and 40 μl of supernatant is mixed with 6 μl of loading buffer and electrophoresed on a 1% agarose gel, stained with ethidium bromide, and viewed with UV illumination.

Discussion

The use of M13KO7 for the production of ss plasmid DNA normally gives titers of cfu of 10^{11}–5×10^{11}/ml and phage titers 10- to 100-fold lower

[17] J. Messing, this series, Vol. 101, p. 20.

(Fig. 4A). Plasmids containing inserts as large as 9 kb have been packaged as ssDNA without a significant loss in yield (M. McMullen and P. Das, personal communication) and instability has not been a problem. It has been observed that some clones, irregardless of size, give reduced levels of ssDNA. This reduction in yield has been both dependent (M. McMullen, personal communication) and independent (J. Braam, personal communication) of the orientation of the insert. M13KO7 has given high yields of ssDNA from pUC-derived vectors, but when it was used as a helper phage with pZ150,[19] a vector constructed from pBR 322, the yield of ssDNA was not significantly different from the yield given by other helper phages. Whether this is due to the lower copy number of pBR as compared to pUC or to some effect of the vector structure is not known. It has been noted that the position and orientation of the IG region within the plasmid can affect its packaging as ssDNA. An example is shown in Fig. 4B (lane 3). This plasmid has the IG region inserted in the same position but the opposite orientation as compared to pUC 119/118, and always gives two bands. However, if the IG region, in the opposite orientation of 118/119, is inserted within the polycloning sites of a pUC vector, the resulting plasmid yields a single band after gel electrophoresis (data not shown). A large variation in the yield of ss plasmid DNA has been seen between different bacterial strains. MV 1184 (derived from JM 83) and MV 1190 (derived from JM 101) have given satisfactory yields. MV 1304 (derived from JM 105) gives much reduced yields and JM 109 undergoes significant lysis when it contains both plasmid and phage.

Acknowledgments

We would like to thank B. McClure, R. Zagursky, M. Berman, and D. Mead for valuable discussions. We also thank M. Volkert for the MV bacterial strains and Claudia Dembinski for help in preparing this manuscript. This work was supported by the Department of Energy, Grant #DE-FG05-85ER13367.

[18] M. Zoller and M. Smith, this series, Vol. 100, p. 468.
[19] R. J. Zagursky and M. L. Berman, *Gene* **27**, 183 (1984).

Fd Cosmet. Toxicol. Vol. 4, pp. 427–432. Pergamon Press 1966. Printed in Great Britain

Review Section

Acceptable Daily Intake of Food Additives and Ceiling on Levels of Use

S. C. HANSEN

The National Health Service of Denmark,
1 St. Kongensgade, Copenhagen, Denmark

(*Received 24 January 1966*)

Introduction

During the last decade many countries have established so-called positive lists of permitted food additives in the light of existing toxicological data as well as the mounting new evidence that has become available. In toxicological evaluations of various classes of food additives, the Joint FAO/WHO Expert Committee on Food Additives (1962, 1964) has introduced the concept of unconditional and condition zones of acceptable daily intake (ADI) of food additives and applied this concept to the compounds under appraisal. Because of the lack of information on food intake, food legislators and administrators have, however, found it difficult to make full use of ADI in the establishment of positive lists with permissible levels of use for food additives. This problem is of current interest since the Joint FAO/WHO Conference on Food Standards (1962) gave first priority to the elaboration of world-wide standards for additives in foodstuffs. The Joint FAO/WHO Codex Alimentarius Commission (1963) and its Committee on Food Additives under the chairmanship of the Dutch Government, is faced with the problem of how to budget for the ADIs so that present and future requirements for additives in foodstuffs can be fairly met, if possible on a world-wide basis.

Regional differences in eating habits are obvious obstacles to the establishment of world-wide standards for permissible levels of use of food additives. It might therefore be fruitful first to consider the diet as a whole and later on to discuss individual foodstuffs. For a few additives ADI is unlimited and therefore presents no difficulties to a world-wide approach. In other words in these few instances ADI poses a ceiling on the permissible levels of use which is infinitely high so that the technological requirements can be met under all circumstances. For other additives ADI is finite, and it would be of interest to investigate whether the ceiling on the allocated permissible levels of use is high enough to permit a world-wide approach. The following considerations aim at a mathematical expression which gives, when the ADI is known, the upper limit of the levels of use which would be acceptable on an international basis.

Calorie intake

Attention is called to some physiological data which, like ADI, can be expressed in relation to body weight, and thus be of assistance in providing criteria for the evaluation of the upper limit for the levels in which additives could be used in foodstuffs.

While many factors influence the intake of the individual foodstuffs and make it difficult to assess the high consumption, the calorie intake follows rather well-known physiological laws (Second FAO Committee on Calorie Requirements, 1957).

The calorie requirements of individuals are determined by four variable factors: (1) physical activity; (2) body size and composition; (3) age; (4) climatic environment.

In assessing the calorie intake it is necessary to consider the variation in the requirements arising under these four headings. The degree of physical activity is certainly the most difficult to assess while the other three variables lend to more precise measurements and the effect of each on calorie requirements can be expressed arithmetically. For the estimation of maximum intake in relation to body weight the variation with age is, however, of paramount interest (Food and Nutrition Board, 1964).

The newborn baby has a calorie requirement of 130 calories/kg body weight. During the first year of life the requirement decreases gradually to 100 calories/kg. A further reduction takes place to 90 calories/kg at the age of four, to 50 calories/kg for the teenagers, and down to 30 calories/kg for a 65-yr-old person. There are in fact wide individual variations in the physical activity of infants and children, and account must also be made of the special energy demands of heavy workers, pregnancy and lactation. A comprehensive survey by Consolazio (1963) of the energy requirements of men living under extreme environmental conditions concludes that:

(1) In extremely cold environments a man performing light physical activities may consume from 35 to 46 calories, for moderate work from 47 to 55 calories, and for heavy physical activities his requirements range from 56 to 68 calories/kg of body weight.

(2) In a temperate environment, the energy requirements for corresponding physical activities are 32–44, 45–52, and 53–63 calories/kg of body weight, respectively.

(3) In extremely hot environments, the energy requirements may range from 55 to 61 calories/kg of body weight for men performing a moderate degree of daily physical activity.

Ceilings

When a child approaches 1 yr of age he partakes of the family meal, although the diet of the very young infant differs somewhat from that of the rest of the family. Bearing in mind that baby food should not contain food additives it seems unlikely that any person will receive more than 100 calories/kg body weight from food carrying additives. It can therefore be assumed without further proof that 100 calories carry less than the ADI. If the ADI is A mg/kg body weight one calorie should carry not more than A/100 mg (Formula 1).

The calorie value of foodstuffs is usually expressed as calories/100 g. Tolerances are often expressed as ppm (parts per million), which is equivalent to mg/kg foodstuff. If a foodstuff provides C calories/100 g, 1 kg would yield 10 C calories and in accordance with formula 1 should not carry more than $(A/100) \times 10$ C mg or

$$T_C = \frac{10 \times A \times C}{100} \text{ ppm (formula 2)}$$

where T_C is the ceiling on the levels of use, based on calorie intake.

Liquid

When liquid intake is expressed in relation to body weight it shows the same variation with age as calorie intake. The requirement of the infant is met under ordinary conditions by a mixture of liquids which together provide approximately 150 ml/100 calories (Food and Nutrition Board, 1964). Galagan, Vermillion, Nevitt, Stadt & Dart (1957) found that fluid intake from infancy through 10 yr of age decreased with time. Their study was carried out over 1 yr in a hot climate and it showed that the total fluid intake at about 2 yr of age had decreased to 1 oz/lb body weight. The standard deviation was 0·4 oz which means that 95% of the children of that age consumed less than 1·8 oz/lb or 120 ml/kg body weight. The milk consumption was included, but not specified for each age group. Table 1 shows the relative proportion of each kind of fluid consumed for all children included in the investigation.

Table 1. *Percentage proportion of total fluids consumed by infants (after Galagan* et al. *1957)*

	Percentage of total fluid intake	
Fluid consumed	Town 1	Town 2
Drinking water	33·8	36·4
Water based beverages	9·6	11·2
Milk	47·9	39·9
Carbonated beverages	3·0	3·6
Other fluids	5·7	8·9

Consolazio (1963) in his survey on the energy requirements of men living under extreme environmental conditions reports that the fluid intake in one strictly controlled study of men working under the extremely hot desert sun averaged 7·9 kg/man/day, i.e. 4·8 kg from beverages, 2·6 from food, and 0·5 kg from oxidation processes in the body.

The above data on liquid intake lead to the conclusion that it is unlikely that any person will ever drink more than 100 ml/kg body weight/day from beverages, excluding milk. In analogy with calories, ceilings on levels of use of food additives might be based on liquid intake. If the ADI is A mg/kg body weight 1 ml should carry not more than A/100 mg, and

$$1000 \text{ ml not more than } \frac{A \times 1000}{100} \text{ mg, or}$$

$$10\,A = T_L \text{ (formula 3)}$$

where T_L is the ceiling on levels of use, expressed in ppm, when it is based on liquid intake.

Table 1 furnishes some evidence that 20 ml/kg body weight/day will be the highest intake of soft drinks. Under this assumption

$$T_L \text{ soft drinks} = \frac{A \times 1000}{20} = 50\,A.$$

ADI allocations to both foodstuffs and beverages

Foodstuffs and beverages may often compete for a share of the ADI and in such cases the ADI must be split into two fractions:

$$ADI = ADI_C + ADI_L$$

where ADI_C is the allocation for foodstuffs and enters into formula 2, while ADI_L is for beverages and enters into formula 3.

T_C and T_L may be raised

T_C and T_L calculated in accordance with formulae 2 and 3 take into account the future possibility that a particular additive may be required in all foodstuffs and beverages throughout the world. To budget for ADI in such a cautious way would have merits because it is difficult to predict which foodstuffs will require which additives. The cautious approach would do no harm if the ceilings are still high enough for the desired technological effect to be accomplished.

It is anticipated that additives will often be required in amounts higher than those indicated by formulae 2 and 3. In such a case the following possibilities should be examined:

(1) Whether the additive in question can be replaced by less toxic substances;

(2) whether the same technological effect can be obtained with smaller amounts in combination with other additives or safe physical treatment;

(3) whether technological reasons indicate that the additive will be used in only a part of the foodstuffs, which would mean that the dividing factor of 100 is substituted by a smaller figure, and T_C or T_L correspondingly raised;

(4) whether it can be decided administratively that the use of this particular additive should be reserved for certain foodstuffs or beverages. In this case T_C or T_L would be raised on basis of information about average food consumption, food consumption trends and possible other limiting factors on the intake of the particular foodstuff or group of foodstuffs.

When the door is kept open for the use of additives in all foodstuffs and beverages, the ceiling on permissible levels of use is found with accuracy from formulae 2 and 3. When, however, in accordance with paras. 3 and 4 above, restricted use is introduced the ceiling cannot be fixed with the same accuracy on a world-wide basis, because the highest intake—on basis of body weight—from the individual foodstuffs is not known, but some basis for an estimation is available as indicated in the following paragraphs.

The Food and Agricultural Organization (1964) tabulated the net food supply in calories per head in 53 countries. Calories were derived from the following ten main groups of foodstuffs: Cereals; potatoes and other starchy foods; sugar and syrups; pulses and nuts; fruit and vegetables; meat; eggs; fish; milk, and milk products excluding butter; fats and oils.

In the affluent countries the consumption has reached a saturation point or ceiling for the following three groups: Sugar and syrups—500 calories/person/day or 16% of calories; milk and milk products excluding butter—400 calories/person/day or 13% of calories, except for a few dairy-producing countries; fats and oils—500 calories/person/day or 16% of calories.

For the groups meat, eggs, and fish there is a clear upward trend, and a saturation point has still to be reached. The highest average consumption from these combined groups is presently approaching 25% of calories. For cereals, potatoes, and other starchy foods there is a clear downward trend in affluent countries.

That the overall consumption of fats and oils has not been affected despite sizeable shifts which have occurred within and between different product groups was confirmed by Hester & Boggs (1964). They pointed out, however, that new applications of modified edible fats and oils will serve to increase the overall consumption.

Application of formulae 2 and 3

Benzoic acid is taken as an example of how permissible levels of use may be established on a world-wide basis. The conditional acceptable daily intake zone is 5–10 mg/kg body weight.

$$\text{Formula 2: } T_C = \frac{A \times C}{10} = \frac{10 \times C}{10} = C \text{ ppm.}$$

Levels of use of benzoic acid are generally permissible if the amount, in ppm, does not exceed the calorie value (kcalories) of 100 g of the foodstuff. On average, the diet (excluding beverages) has a calorie value of approximately 200 calories/100 g. Even in combination with other preservatives or with physical treatment 200 ppm benzoic acid would be rather low. The possibility should be investigated whether the use of benzoic acid can be reserved for a certain part of the diet and the ceiling be raised accordingly. It should be fully realized that if this does happen it will almost be impossible to widen the use and lower the ceiling later on, because meanwhile standards may have been established up to the higher ceiling.

Sodium benzoate and benzoic acid are very widely used (Food and Nutrition Board, 1956), for example in the following products with levels of use (%) given in parentheses: margarine, mayonnaise (0·1); semi-preserved fish products (0·1); incorporated in ice for cooking fish (–); maraschino cherries (0·1); mincemeat (–); bottled soft drinks (0·05); fruit juices (0·1); pickles (0·1); confections (0·1); marmalade (0·1).

The largest part of calories derived from the above foodstuffs belongs to the groups "Fats and oils" and "Sugar", which each contribute on average 16% of calories. Some calories come from fish and a minor contribution is derived from fruits and vegetables. To obtain a significant rise of the ceiling the use of benzoic acid may be reserved for semi-preserved fish, beverages, and for those foodstuffs where the main part of calories are from sugar. The total intake of calories from these foodstuffs are not likely to exceed 20 calories/kg body weight. ADI has to be split into two fractions, e.g.

$$ADI = ADI_C + ADI_L$$

$$10 \text{ mg} = 5 \text{ mg} + 5 \text{ mg}$$

$$\text{Formula 2 gives: } T_C \text{ benzoic acid} = \frac{5 \times C}{10} \times \frac{100}{20} = 2{\cdot}5\ C \text{ ppm.}$$

$$\text{Formula 3 gives: } T_L \text{ benzoic acid} = 50 \times 5 = 250 \text{ ppm.}$$

For a foodstuff with a calorie value corresponding to that of sugar (400 calories/100 g) a level of use up to a maximum of 1000 ppm is permissible. Beverages may contain up to 250 ppm, but if they have a significant calorie value due to added sugar, e.g. 100 calories/100 g, they may contain more, in this example $250 + 2{\cdot}5 \times 100 = 500$ ppm.

GENERAL CONCLUSIONS

To assist the food legislator and administrator in his work on the establishment of permissible levels of use for additives (tolerances) it is suggested that the acceptable daily intake zones given by toxicologists should be converted into ceilings on levels of use. Modern food processing and marketing demand the use of additives in a variety of foodstuffs, and it is very difficult to predict future technological needs, but calculated on basis of high calorie and liquid intake per kg body weight per day the ceilings will function as a budgeting system ensuring that any request for additives below the ceiling can be met now and in the future.

When additives are required in amounts higher than indicated by the ceiling the possibili-

ties should thoroughly be investigated whether the additive in question could be replaced by a less toxic substance, or whether the same technological effect could be obtained with smaller amounts in combination with other additives or safe physical treatment. It may, however, be often necessary to raise the ceiling by reserving the use of an additive for certain foodstuffs of which high intake must be estimated in terms of calories or ml per kg body weight.

SUMMARY

Administratively fixed levels of use for food additives (tolerances) should accord with the acceptable daily intake (ADI) zones established by toxicologists. ADI is expressed in mg/kg body weight. The highest intake of calories is estimated to be 100 calories/kg body weight/day (which is that of a 1-yr-old baby). It can therefore be assumed without further proof that all foodstuffs contain less than ADI/100 calories. If higher levels of use are required, the additive in question must be reserved for specific foodstuffs or groups of foodstuffs from which the highest intake, as calories/kg body weight/day, must be estimated. ADI may also be related to liquid intake.

Acknowledgements—I wish to express my gratitude to my colleagues in the National Health Service, in the Danish National Codex Alimentarius Committee, and to former colleagues in the FAO and WHO for valuable suggestions.

REFERENCES

Consolazio, C. F. (1963). The energy requirements of men living under extreme environmental conditions. *World Review of Nutrition and Dietetics*. Vol. 4. Edited by G. H. Bourne. p. 55. Pitman Medical Publishing Co. Ltd., London.

Food and Agricultural Organization (1964). *Production Yearbook*. Vol. **18,** p. 259.

Food and Nutrition Board (1956). *The Use of Chemical Additives in Food Processing*. Publication 398. National Academy of Sciences–National Research Council, Washington, D.C.

Food and Nutrition Board (1964). *Recommended Dietary Allowances*. 6th revised ed. Publication 1146. National Academy of Sciences–National Research Council, Washington D.C.

Galagan, D. J., Vermillion, J. R., Nevitt, G. A., Stadt, Z. M. & Dart, Ruth E. (1957). Climate and fluid intake. *Publ. Hlth Rep., Wash.* **72,** 484.

Hester, O. C. & Boggs, R. L. (1964). *Market Potentials for Modified Edible Fats & Oils*. US Department of Agriculture, Economic Research Service, Marketing Economics Division. Marketing Research Report No. 659, Washington D.C.

Joint FAO/WHO Codex Alimentarius Commission (1963). Report of First Session, Rome.

Joint FAO/WHO Conference on Food Standards (1962). Report on Food Standards, Geneva, p. 15.

Joint FAO/WHO Expert Committee on Food Additives—Sixth Report (1962). Evaluation of the Toxicity of a Number of Antimicrobials and Antioxidants. *Tech. Rep. Ser. Wld Hlth Org.* **228.**

Joint FAO/WHO Expert Committee on Food Additives—Seventh Report (1964). Specifications for the Identity and Purity of Food Additives and their Toxicological Evaluation: Emulsifiers, Stabilizers, Bleaching and Maturing Agents. *Tech. Rep. Ser. Wld Hlth Org.* **281.**

Second FAO Committee on Calorie Requirements (1957). *Calorie Requirements. FAO Nutritional Studies* No. 15, Rome.

Food Additives and Contaminants, 1997, Vol. 14, No. 8, 791–802

Evaluation of the Budget Method for screening food additive intakes

Judith S. Douglass†, Leila M. Barraj†, David R. Tennant‡§, Wesley R. Long† and Christine F. Chaisson†

† *Technical Assessment Systems, Inc., Washington DC 20007, USA*
‡ *TAS International, Malvern WR14 2AZ, UK*

(Received 4 April 1997; accepted 8 April 1997)

The Budget Method, originally developed for determining food additive use limits, has been proposed as a tool for screening food additive intakes to establish monitoring priorities. Theoretical Maximum Daily Intake (TMDI) estimates derived using the Budget Method rely on assumptions regarding physiological requirements for energy and liquid and on the energy density of food rather than on food consumption survey data. This report summarizes work performed to determine the validity of Budget Method assumptions and to assess the potential for error in assigning monitoring priority based on Budget Method results. Budget Method assumptions regarding energy and liquid intake were compared with data from UK, German and US nationwide food consumption surveys. It was found that the Budget Method assumptions of energy intake and liquid intake are higher than mean intakes reported in surveys. The Budget Method assumption regarding energy density of foods also was found to be a slight overestimate. Budget Method TMDIs for case study additives were in each case larger than survey-based 95th percentile per capita *additive intake estimates. Based on these results, the Budget Method appears to be a suitably conservative screen for establishing additive monitoring priorities based on potential lifetime average intakes.*

Keywords: Food additives, intake, exposure, energy intake, risk assessment, monitoring

§To whom correspondence should be addressed.

Introduction

European Union (EU) Directives 94/35/EC, 94/36/EC and 95/2/EC require each Member State to monitor the consumption and usage of food additives. The European Commission was not empowered by these directives to specify how the Member States should fulfil this obligation. However, the Commission does have a role in co-ordinating the national programmes and is called upon to submit a report to the European Parliament. The Commission recognizes that a high degree of co-ordination is essential if its report to the European Parliament is to provide a coherent summary of the overall situation in the European Union.

A variety of additive monitoring methods has been discussed. A mega-database of information on additive concentrations in foods has been proposed but such a database is unlikely to be established in time to meet the requirements of the EU Directives (Nutriscan 1994). *Ad hoc* chemical analyses might be appropriate in specific applications, but would be impractical for effective monitoring of all chemicals. Targeted surveys of use and/or intake of specific additives can be (and have been) performed to provide important data for monitoring but the cost of monitoring the use of all additives in all foods by chemical analysis would be prohibitively expensive.

Many of the methods in use for determining intakes of additives rely on food production statistics or on national food consumption survey data. In general, estimates of additive intake from food consumption data must be generated by experts who understand the nuances and limitations of survey techniques and know the details of the particular survey. The most accurate estimates are generated using data on consumption of specific foods by individuals and such data are not widely available. Additive intake estimation using food consumption data, while requiring fewer resources than the analytical approaches discussed above, still requires significant

0265-203X/97 $12·00

resources and is considered impractical for use with every additive.

Although Member States are not required to use a common monitoring approach, the development of a science-based but practical monitoring system is a formidable task which they all face. Eleven Member States and Norway, acting under the EU Scientific Co-operation Directive (SCOOP), have been working together to develop a tiered system for prioritizing needs for monitoring specific additives. The Codex Alimentarius Commission Committee on Food Additives and Contaminants has considered such a system (CX/FAC/96/6) for determining whether the additive uses listed in the Codex General Standard for Food Additives (GSFA) pose any risk to public safety.

The 'tiers' described in the Codex protocol are essentially additive intake estimation methods which progress in complexity and data requirements. These methods are intended to produce progressively more accurate estimates of additive intake, although they have not been validated for this purpose. Where results of analyses in a low level tier indicate that an ADI is unlikely ever to be exceeded, the screen eliminates the additive from further consideration, thus conserving valuable resources. Resources can then be focused toward additives for which the potential for exceeding the ADI cannot be excluded with the present conditions of use and assumptions used in the assessment. In these cases, additives are further evaluated using more complex but less conservative intake estimation methods.

The lower level tiers are designed to be ultra-conservative in order to minimize the risk that an additive, for which there is a possibility that the intake may exceed the ADI, is erroneously eliminated from further consideration. However, it must be emphasized that these tiers are tools for establishing monitoring priorities rather than tools for risk assessment; results from screening tests should never be interpreted as realistic estimates of intake.

An adaptation of the Budget Method, developed by Søren Hansen (1966) of the National Food Agency of Denmark for determining maximum use levels for food additives, has been proposed by Codex as the Tier 1 screening step. Use of the Budget Method in the manner proposed by Codex is intended to yield a worst-case additive intake estimate, the 'theoretical maximum daily intake' (TMDI).

The Budget Method provides the basis for a simple, inexpensive screening method for intakes because it relies on assumptions regarding physiological requirements for energy and liquid and on energy density of food rather than on food consumption survey data. The TMDI is calculated by assuming that all foods contributing to the energy intake, and all beverages contributing to the liquid intake, will contain the additive at maximum permitted use levels. Under the Codex proposal, an additive is said to 'pass' the Budget Method screen if the calculated TMDI is lower than the additive's acceptable daily intake (ADI).

The Budget Method may also be an appropriate screening method for determining priorities for additive monitoring required under EU Directives 94/35/EC, 94/36/EC and 95/2/EC. Before it can be accepted in the EU for such a purpose, however, the underlying assumptions must be examined to evaluate the potential for error in the use of results to establishing monitoring priorities. Clearly, the Budget Method must *over*-estimate intake in order to minimize Type II (false negative) errors; a false indication could result in unchecked use of a substance which should have been a priority for monitoring. Conversely, Type I (false positive) errors should be minimized to prevent unnecessary expenditure of time and resources in pursuit of more detailed intake estimates.

This report summarizes work performed by TAS International, at the request of the International Life Sciences Institute Europe (ILSI Europe), to evaluate the validity of assumptions on which the Budget Method is based and to assess the potential for Type I and Type II errors in using results to establish additive monitoring priority.

The Budget Method

The Budget Method was designed to convert food additive ADIs into 'ceilings of use' calculated on the basis of maximum intakes of food and beverages potentially containing the additives (Hansen 1966, 1979). In budget calculations for additives used in both solid foods and beverages, the ADIs are divided into two fractions. The proportion of the ADI allocated to food and the proportion allocated to beverages are decided upon arbitrarily to accommodate technological requirements.

In developing the Budget Method for determining ceilings for additives used in foods, Hansen started from a basic assumption that from the time a child begins eating food potentially containing additives, energy intake will not exceed 100 kcal/kg body weight/day. He supported this assumption with data showing an inverse relationship between age and energy requirements per unit body weight, noting that 'there is a remarkable steep decline during the first few years of life which enables us to identify a landmark or starting point for estimation of intake, namely 100 kcal per kg of body weight per day'. Hansen used a conversion factor of 2 kcal/g for 'average food including milk but excluding other beverages' to estimate the maximum amount of foods containing an additive as 50 g food/kg body weight/day. Additive ceilings can then be calculated as being equal to *one ADI per 50 g of food or (ADI × 20) mg/kg food.*

Hansen recognized that by assuming an intake of 50 g food/kg body weight/day, with all food potentially containing all additives, additive ceilings calculated using the Budget Method could fall below technologically effective levels. Believing that the ADI safety factors cover differences between adults and children, Hansen stated that he was 'unduly cautious' in basing calculations on very young children, and judged that 'a factor of 2 should be permissible' for adjusting food additive ceilings to *(ADI × 40) mg/kg food.*

Hansen also proposed adjustment of additive ceilings to technologically effective levels, where necessary, by restricting use of relevant additives to half or a quarter of the 'full horizon' of foods, using assumptions related to proportions of processed food in the food supply. These assumptions have been interpreted as meaning that processed food does not represent more than 50% of total maximum food intake and that half of this is represented by processed milk, meat, fish, poultry, vegetables, and cereal products less likely to contain additives (Hallas-Møller 1987, Bär and Würtzen 1990). Food additive ceilings would then be *(ADI × 80) mg/kg food or (ADI × 160) mg/kg food*, depending on whether the predicted additive use pattern included processed milk, meat, fish, poultry, vegetables and cereal products. The conversion from 'additive ceiling' to TMDI and the factors involved are described in table 1.

In determining ceilings of use for additives used in beverages, Hansen reviewed the literature on liquid intakes of infants, children, and adults in hot climates and concluded that 'it is unlikely that any person will ever drink more than 100 ml/kg body weight/day from beverages, excluding milk'. Hansen also referred to data on liquid requirements, stating that 'if we choose the recommended liquid intake at the age of two, 100 ml per kg body weight per day, as the basis for our calculations of intake we cover the child and we also cover the adult'. Hansen proposed adjustment of additive ceilings to technologically effective levels, where necessary, by applying factors to double or redouble the ceilings, considering competition among beverages. In assessing additive ceilings for soft drinks, the assumption has been that soft drinks account for 25% of the 100 ml/kg body weight/day maximum intake (Hallas-Møller 1987, Bär and

Table 1. Assumptions made in screening additive intakes using the Budget Method: additives used in food.

1 Maximum energy intake by young children consuming table foods is 100 kcal/kg body weight/day.
2 Maximum energy intake over the course of a lifetime is 50 kcal/kg body weight/day.
3 50 kcal = 25 g food.

If:	*Then the Theoretical Maximum Daily Intake (TMDI) of an additive can be calculated as:*
4 The additive is used in foods other than baby foods, and the maximum amount of food required to meet energy requirements is: 25 g per kg body weight	$\frac{\text{Maximum use level (mg additive/kg food)}}{40}$ (NB: 25 g food = 1/40 kg food)
5 50% of foods consumed are processed.	$\frac{\text{Maximum use level (mg additive/kg food)}}{80}$
6 The additive is used in processed foods other than those considered to be important in the daily diet (e.g. dairy, meat, fish, poultry, vegetable or cereal products).	$\frac{\text{Maximum use level (mg additive/kg food)}}{160}$

Table 2. Assumptions made in screening additive intakes using the Budget Method: additives used in beverages.

	If:	*Then the Theoretical Maximum Daily Intake (TMDI) of an additive can be calculated as:*
1	Maximum intake of liquids (other than milk) is 100 ml/kg body weight/day.	
2	The additive is used in all non-milk beverages, and the maximum amount of non-milk beverages consumed is $\frac{100\text{ml}}{\text{kg body weight}}$ (100 ml = 1/10 litre).	$\frac{\text{Maximum use level (mg additive/l beverage)}}{10}$
3	The additive is used only in soft drinks, and maximum soft drink intake is 25% of non-milk beverage intake.	$\frac{\text{Maximum use level (mg additive/l beverage)}}{40}$

Würtzen 1990). Assuming a density of 1 g/ml, the ceiling for additive use in soft drinks can be calculated as *(ADI × 40) mg/kg beverage*. The conversion from the beverage 'additive ceiling' to TMDI and the factors involved are described in table 2.

The validity of Budget Method assumptions for calculating additive ceilings of use has been evaluated on a limited basis by Hallas-Møller (1987) using the results of a 1985 Danish food consumption survey of adults (National Food Agency of Denmark 1986). Assumptions regarding maximum food intake (50 g/kg body weight/day) and the proportion of processed food consumed (50%) were judged valid for 90% of adults. However, up to 25% of adults consumed more 'high additive' foods than the 25% of total foods assumed by Hansen.

Hallas-Møller found the Budget Method assumption regarding maximum liquid intake (100 ml/kg body weight/day) to be valid for 99·6% of Danish adults. Bär and Würtzen (1990) reviewed liquid intake data from the US and several European countries, and concluded that 'a daily liquid consumption of 100 ml/kg body weight is not a reasonable starting point for budget calculations. Instead it appears that a daily liquid intake of 50 ml/kg will cover the needs under extreme conditions, and of heavy consumers'.

Bär and Würtzen (1990) examined the value and limitations of the Budget Method for calculating additive use ceilings using intense sweeteners as examples. The authors found that intense sweetener intakes estimated using data from targeted surveys of sweetener intake by several European populations were far below intakes predicted by the Budget Method. The authors proposed several modifications including alternate reference points and correction factors to compensate for competing additives with similar functional uses. They noted a potential need for additional correction factors to be used with additives applied in limited types of foods and with additives which have self-limiting properties.

The proposal to use the Budget Method as Tier 1 of the Codex GSFA safety evaluation calls for calculation of TMDIs based on maximum current additive use levels. Assumptions made in TMDI calculations are listed in tables 1 and 2. Use of the Budget Method in EU screening to establish monitoring priority would, if accepted for this purpose, involve similar assumptions and calculations. However, the Budget Method assumptions to be used in these calculations were developed for purposes other than those proposed, and must be carefully evaluated for relevance to the proposed applications.

Methods

Evaluation of underlying assumptions

To evaluate whether Budget Method assumptions regarding energy intake, energy density of food, beverage consumption and soft drink consumption provide a valid basis for screening additive intakes to determine priority for monitoring, TAS examined intake data from nation-wide food consumption surveys of individuals conducted in the United Kingdom, the former West Germany, and the United States (table 3). The Budget Method assumes that ADIs for additives relate to average intakes over a

Table 3. Food consumption survey data used in Budget Method validation.

United Kingdom
Data used in the analyses were taken from the following summary sources of survey data published by the Ministry of Agriculture, Fisheries and Food:

- *Food and Nutrient Intakes of British Infants Aged 6–12 Months*; 1986; 488 infants; based on 7-day food intake records.
- *National Diet and Nutrition Survey: Children Aged $1\frac{1}{2}$–$4\frac{1}{2}$ years;* July 1992 and June 1993; 1675 children; based on 7-day food intake records.
- *The Diets of British School Children;* 1983; 3581 children ages 10–11 and 14–15; based on 7-day food intake records.
- *The Dietary and Nutritional Survey of British Adults*; October 1986–August 1987; 2197 adults aged 16 to 64; based on 7-day food intake records.

Former West Germany
National Consumption Study (NVS); October 1985–January 1989. Seven-day weighed intake data were collected from over 25 000 individuals 4 years of age and older, using a system consisting of over 6000 food codes. Data used in the present analyses were taken from a dataset containing records of average daily intakes of foods in 90 summary food groups by individual survey respondents.

United States
US Department of Agriculture 1989–90, 1990–91 and 1991–92 Continuing Surveys of Food Consumption by Individuals (CSFII). Together, the CSFII surveys measured dietary intake of over 11 000 individuals over a 3-day period. Although these data clearly are not directly reflective of European food intake patterns, it was felt that the extensive information on intake by individuals could be of potential value in examining basic Budget Method assumptions regarding food additive intakes. For example, neither the UK surveys nor the German survey collected data on individuals in all age groups; the US data provide supplementary data on total population intakes.

liketime. Additive intakes are unlikely to be sustained at high levels throughout life. The validity of the underlying assumptions was therefore assessed on the basis of population averages rather than on high level intakes in order to be consistent with the Budget Method approach.

Budget Method assumptions regarding proportions of the diet accounted for by processed foods potentially containing additives were not tested due to limitations in the availability of European survey data. The UK surveys were conducted on selected age groups, and results available for analysis were summarized for age groups and broad food groups. The West German survey results used were summarized by food group, but represented individual consumption by all respondents in the survey population (4 years of age and older). The US results used represented consumption of specific foods by each of the survey respondents. Analyses were conducted on four representative age bands corresponding to the UK food consumption survey data.

Case studies

To evaluate the potential for Type I and Type II errors in use of the Budget Method as a screening tool, case studies were conducted for two food additives with different characteristics and proposed uses (tables 4 and 5). The additives are hypothetical. However, maximum permitted use levels in specific food groups were selected to reflect realistic use levels. Sample Food Additive 1 was designed to be representative of an additive used at consistent levels in a broad range of foods consumed by a significant portion of the population (e.g. a stabilizer or preservative). Food Additive 2 was designed to be representative of an additive used in varying

Table 4. Case Study Additive 1: use limits.

	Additive use limits	
	Food category	Use limit (mg/kg food)
Additive 1: Used at consistent levels in a broad range of foods consumed by a significant portion of the population	Breads	50
	Baked products	50
	Pastas	30
	Cereals	50
	Rice	30

Table 5. Case Study Additive 2: use limits.

	Additive use limits	
	Food category	Use limit (mg/kg food or mg/l beverage)
Additive 2: Used in varying concentrations in a range of foods consumed by specific segments of the population	Soft drinks	350
	Biscuits	1000
	Confectionery	500

Table 6. Budget Method TMDI estimates for Case Study Additives 1 and 2.

Food additive	Use level in calculation	Budget method factor	TMDI (mg/kg body weight)
1	50 mg/kg	80	0·63
2	350 mg/l	40	8·75 } 15·00
	1000 mg/kg	160[a]	6·25 }

[a] Approach recommended by Hallas-Møller (1995)

concentrations in a range of foods consumed by specific segments of the population (e.g. an intense sweetener).

The Budget Method TMDI for Additive 1 was calculated assuming that up to 50% of consumed foods would contain the additive, using the food intake factor of 80 derived in the previous section (table 6). The TMDI for Additive 2 was calculated allocating half of the budget to soft drinks and half to 'high additive' solid foods (biscuits and confectionery), using a liquid intake factor of 40 and a food consumption factor of 160.

Food consumption survey-based intake estimates for Additives 1 and 2 were calculated as precisely as possible given the particular limitations of the surveys. Additive use levels were applied to food consumption data in the categories shown in tables 7 and 8.

The Budget Method is generally assumed to provide a conservative estimate of the upper limit of lifetime intake for all consumers. In order to test this, high level (95th percentile) intakes of adults (16–64 years) were used for comparison with Budget method TMDIs.

Limitations of the survey data used in these analyses precluded calculation of per-user intakes. However, because it was assumed that all people will be consumers of the additives in question at some point in their lifetimes, *per capita* estimates were judged to be representative of lifetime intakes. It was also assumed that extreme high-level user intakes on the per kg body weight basis would not be maintained throughout the course of a lifetime, and that 95th percentile *per capita* intakes would provide a reasonable picture of upper level intake. Where available survey data did not permit assessment of 95th percentile *per capita* intakes, these intakes were estimated using the Bernier method (Bernier *et al.* 1994), by taking three times the mean.

Results

Validation of Budget Method assumptions regarding energy intake, energy density of foods, beverage consumption and soft drink consumption

Energy intake. Budget Method assumptions regarding maximum energy intake are 100 kcal/kg body weight/day for 1-year old children and 50 kcal/kg body weight/day over a lifetime.

Mean energy intakes reported in the UK, German and US surveys are shown in table 9. The energy intake distribution for the total population and for 1-year olds, based on US survey data, is presented in figure 1.

Energy intakes reported in the UK, German and US surveys are comparable across populations; the results confirm that as with energy requirements, daily energy intake, adjusted by body weight, decreases with age. The mean energy intake for infants falls within the upper tail for energy intake for the total population. The survey results therefore indicate, at least in the populations examined, that the Budget Method assumptions of 100 kcal/kg body weight energy intake for 1-year olds and of 50 kcal/kg body weight for the general population tend to overestimate actual lifetime intakes of 91–96 kcal/kg body weight for 1-year olds and 35–39 kcal/kg body weight for the general population.

The extent of the apparent Budget Method overestimate of intake may not be significant, as under-reported energy intake has been documented in research on various survey techniques (Bingham 1987, Schoeller 1990, Black *et al.* 1993, Haroldsdottir *et al.* 1993). Using a low ratio of energy intake to estimated basal metabolic rate as the criterion for under-reported energy intake, Briefel *et al.* (1995) found under-reporting in two major US surveys, one of which was the CSFII used in analyses for the present report.

Energy density of food. In developing the Budget Method, Hansen (1966, 1979) assumed that 50 grams of food had an energy value equal to 100 kcal,

Table 7. Case Study Additive 1: additive use categories and corresponding food categories in survey data summaries.

	Food categories in survey data summaries					
	UK					
Additive use categories	*Food and Nutrient Intakes of British Infants Aged 6–12 Months*, 1986	*National Diet and Nutrition Survey: Children Aged 1½–4½ years*, July 1992 and June 1993	*The Diets of British School Children*, 1983	*The Dietary and Nutritional Survey of British Adults*, October 1986–August 1987	West German NVS, 1985–89	US CSFII, 1989–92 (detailed food codes within the following categories)
Breads	Bread	White bread Wholemeal bread Soft grain bread Other bread	Total bread	White bread Wholemeal bread Other bread	Knäckebrot Vollkornbrot und-brötchen Weissbrot sonstiges Brot	Yeast breads, rolls
Baked products	Biscuits and crispbread Cakes, buns and puddings	Biscuits Fruit pies Buns, cakes and pastries	Bran products Buns and pastries Cakes Biscuits	Biscuits Fruit pies Buns, cakes, and pastries	Kleingebäck Feingebäck, Dauerbackwaren	Cookies
Pastas	Pasta and rice	Pasta	Pasta	Pasta	Teigwaren	Pastas
Cereals	Breakfast cereals	Wholegrain and high fibre breakfast cereals Other breakfast cereals	Breakfast cereals	High fibre breakfast cereals Other breakfast cereals Other cereals	Weizenmehl und-griess Getreidekörner Erzeugnisse aus Getreide sonstige stärkehaltige Produkte	Cooked cereals Ready-to-eat cereals
Rice	(Accounted for in Pasta and rice category, as indicated above)	Rice	Rice	Rice	Reis	Rice

Table 8. Case Study Additive 2: additive use categories and corresponding food categories in survey data summaries.

	Food categories in survey data summaries					
	UK					
Additive use categories	*Food and Nutrient Intakes of British Infants Aged 6–12 Months*, 1986	*National Diet and Nutrition Survey: Children Aged 1½–4½ years*, July 1992 and June 1993	*The Diets of British School Children*, 1983	*The Dietary and Nutritional Survey of British Adults*, October 1986–August 1987	West German NVS, 1985–89	US CSFII, 1989–92 (detailed food codes within the following categories)
Soft drinks	Squash and soft drinks	Diet soft drinks Other soft drinks	Colas Fizzy drinks Other soft drinks	Diet soft drinks Other soft drinks	Limonaden und sonstige Saftgetränke	Soft drinks
Biscuits	Biscuits and crispbread	Biscuits	Biscuits	Biscuits	Kleingebäck	Cookies
Confectionery	Chocolate confectionery Sugar confectionery	Sugar confectionery Chocolate confectionery	Chocolate Sweets	Sugar confectionery Chocolate confectionery	Schokolade Schokoladenerzeugnisse Zuckerwaren	Candies

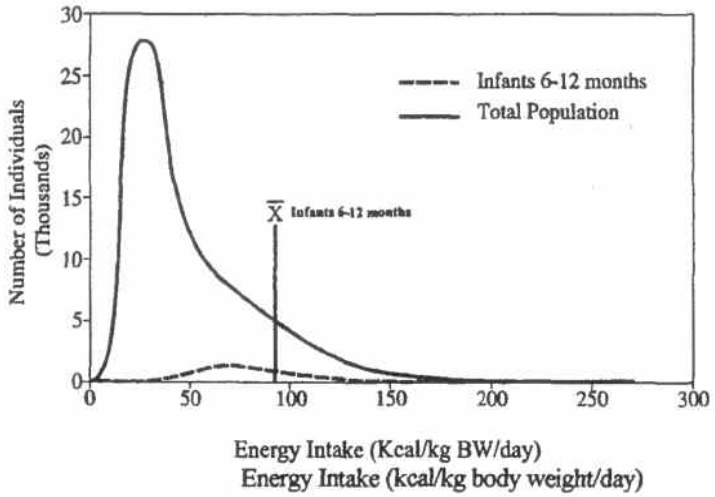


Figure 1. Mean daily energy intake distribution for the total population and for 1-year-olds, based on US survey data.

Table 9. Mean per capita *daily energy intake based on UK, West German and US survey data.*

Age group	Mean *per capita* energy intake (kcal/kg body weight/day)		
	UK	West Germany	US
Total population	NA	39[a]	35
6–12 months	96	NA[b]	91
$1\frac{1}{2}$–$4\frac{1}{2}$ years	76	NA	84[c]
10–11 years	44	59	50
14–15 years	49	44	36
16–64 years	34	35	26

[a] Ages 4+.
[b] NA, Not applicable (population group not included in survey).
[c] Ages 1–5 years.

Table 10. Energy density of food based on UK and US survey data.

Population	Energy density of food consumed, including milk products but excluding other beverages (kcal/g)
UK (ages 16–64)	1·63
US (total population	1·41

which is equivalent to two calories per gram. In contrast, UK and US survey data (table 10) indicate that the mix of food consumed, on average, is of slightly lower energy density (the collapsed German data used in this analysis did not permit calculation of energy density).

It appears that the Budget Method is based on an overestimate of energy density of foods consumed by the general population. Presumably, the extent of the overestimate would be even greater for sub-populations consuming large quantities of low calorie foods. However, the discrepancy between the energy density calculated here and that calculated by Hansen may be due to differing interpretations of what constitutes 'food' and what constitutes 'beverages'. For example, soup was considered a food in the present study, but may have not been considered so by Hansen.

Beverage and soft drink intake. Budget Method calculations for additive intakes from beverages are based upon the assumption that 100 ml/kg body weight of fluids, excluding milk, represents maximum consumption. The maximum soft drink consumption is assumed to be 25% of this consumption, or 25 ml/kg body weight/day.

Survey data on beverages are recorded in grams rather than millilitres, but can be compared in a general way if it is assumed that most beverages are as dense or denser than water, and have a density of 1 g/ml. Beverage consumption data are shown in table 11. It is likely that UK soft drink intakes by $1\frac{1}{2}$–$4\frac{1}{2}$-year-old children are higher than those by children of comparable ages in Germany and the US because water-diluted rose hip, blackcurrant, and other fruit cordials popular with young children in the UK are included in summarized soft drink intake data.

Table 11. Beverage and soft drink intake estimates based on UK, West German and US survey data.

	Mean *per capita* intake of non-milk beverages and of soft drinks (g/kg body weight/day)					
	UK		West Germany		US	
Age group	All non-milk beverages	Soft drinks only	All non-milk beverages	Soft drinks only	All non-milk beverages	Soft drinks only
Total population	NA[a]	NA	17[b]	2[b]	13	4
6–12 months	18	4	NA	NA	17	2
1½–4½ years	31	25[c]	NA	NA	18[d]	5[d]
10–11 years	8	4	16	5	11	5
14–15 years	9	3	13	4	11	6
16–64 years	23	2	18	2	13	4

[a] NA, Not applicable (population group not included in survey).
[b] Ages 4+.
[c] Water-diluted rose hip, blackcurrant, and other fruit cordials popular with young children in the UK are included in summarized soft drink intake data.
[d] Ages 1–5 years.

Table 12. Budget Method TMDI and survey-based intake estimates for Case Study Additives 1 and 2.

		Intake estimates based on food consumption survey data			
Food additive	Budget Method TMDI (mg/kg body weight)	Population	95th percentile *per capita* intake (mg/kg body weight/day)	Percent consuming foods containing additive	Percent total energy intake from foods containing additive (consumers)
1	0·63	UK[a]	0·34	90	33
		West German[b]	0·40	100	27
		US	0·43	99	25
2	15·00	UK	2·88	80	9
		West German	4·42	96	5
		US	5·12	78	8

[a] Ages 16–64; 95th percentile intake for this population computed as (mean intake × 3), as recommended by Bernier *et al.* (1994).
[b] Ages 4+.

Comparison of TMDIs with additive intake estimates based on food consumption survey data

Budget Method TMDIs calculated for Food Additives 1 and 2 (0·63 and 15·00 mg/kg body weight) are above the 95th percentile *per capita* estimates of intake for adults from all three surveys (table 12). However, in these case studies the proportion of the population consuming foods containing the additives is high (78–100%). If smaller proportions of the population were consuming foods containing the additive then *per capita* estimates of intake could under-estimate true levels of intake and the Budget Method would then be less conservative.

The Budget Method TMDI calculated for Food Additive 1 is within a factor of 2 of the survey 95th percentile values, indicating a close, yet still conservative estimate of intake. It is likely that no Type I or

Type II errors would occur in assignment of monitoring priority for Food Additive 1 using the Budget Method.

The scenario for Food Additive 2 is very different. The Budget Method TMDI for Food Additive 2 is 3–5 times the survey-based estimates of intake, indicating a potential for Type I (false positive) error in screening to establish monitoring priority.

Additives 1 and 2 are each used in a wide variety of foods, and a large proportion of the population consumes one or more foods in the contributing food groups. It is therefore difficult to evaluate the relative importance of factors contributing to differences in the extent to which Budget Method assessments might be useful for screening seemingly similar additives. Contributing factors may include relative energy contribution of solid foods containing the additives and relative volume contribution of beverages containing the additives.

Relative energy contributions of foods containing additives. Food Additive 1 is used in foods that contribute 25–33% of the total energy intake.The relative contribution of the foods in which Food Additive 2 may be used is much smaller, ranging from 5 to 9%. The Budget Method assumes that for solid foods, at least 25% of total food energy (using the factor of 160) and perhaps more often 50% of total food energy (using the factor of 80) will be contributed by foods containing the additive. The extent to which the Budget Method TMDI overestimates intake of additives present in foods contributing less than 25% of total energy compared with intake of additives contributing 25% of total energy or more should be tested in future studies.

Volume of liquids containing additives. The TMDI for Food Additive 2 is based on use in soft drinks as well as in solid foods. The assumption in the Budget Method is that soft drink consumption is 25 ml/kg body weight/day. However, as demonstrated in table 11, soft drink consumption is below the assumed level in each survey total population. The impact of the difference between Budget Method assumptions about soft drink consumption and actual population soft drink consumption on intakes of additives such as Food Additive 2 should be investigated further.

Conclusions

Budget Method assumptions regarding energy intake, energy density of foods, beverage consumption and soft drink consumption suggested the following conclusions:

- Assumptions regarding energy intake, beverage consumption and soft drink consumption of the general population are overestimates of actual average levels.
- The Budget Method assumption regarding energy density of foods may be an overestimate.

Budget Method TMDIs were in each of the two theoetical cases studied larger than survey-based *per capita* additive intake estimates for adults even when high-level intakes were considered. This analysis provides evidence that the Budget Method produces conservative estimates of average additive intakes over a lifetime. In fact, potential for Type I (false positive) errors in assignment of monitoring priority was observed in one of the case studies. The evidence provided is limited, but logic predicts that the method will tend to be conservative, and the potential for Type II (false negative) error is judged to be small.

Thus the Budget Method appears to be a conservative first screen for establishing priorities for monitoring consumption and use of food additives based on potential lifetime average exposures. The Budget Method may not be suitable for additives where there may be concerns about exposures over periods of less than a lifetime.

Acknowledgements

The authors gratefully acknowledge the support of the ILSI Europe Food Chemical Intakes Task Force.

References

BÄR, A., and WÜRTZEN, G., 1990, Assessing the use of additives in food: a reappraisal of the Danish Budget Method. *Lebensmittel-Wissenshaft und-Technologie*, **23**, 193–202.

BERNIER, J. J., CAUSERET, J., HOELLINGER, H., and SUSCHETET, M., 1994, Méthode pour assurer, au sein d'une population, le non-dépassement de la dose journalière admissible des additifs alimentaires. *Medicine et Nutricion*, **1**, 9–17.

BINGHAM, S. A., 1987, The dietary assessment of individuals: methods, accuracy, new techniques and recommendations. *Nutrition Abstracts and Reviews*, **57**, 705–742.

BLACK, A. E., GOLDBERG, G. R., JEBB, S. A., LIVINGSTONE, M. B. E., COLE, T. J., and PRENTICE, A. M., 1991, Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys. *European Journal of Clinical Nutrition*, **45**, 583–599.

BLACK, A. E., PRENTICE, A. M., GOLDBERG, G. R., JEBB, S. A., BINGHAM, S. A., LIVINGSTONE, M. B. E., and COWARD, W. A., 1993, Measurements of total energy expenditure provide insights into the validity of dietary measurements of energy intake. *Journal of the American Dietetic Association*, **93**, 575–579.

BRIEFEL, R. R., MCDOWELL, M. A., ALAIMO, K., CAUGHMAN, C. R., BISCHOF, A. L., CARROLL, M. D., and JOHNSON, C. L., 1995, Total energy intake of the US population; the third National Health and Nutrition Examination Survey, 1988–1991. *American Journal of Clinical Nutrition*, **62** (suppl), 1072S–1080S.

HALLAS-MØLLER, T., 1987, Calculating possible daily intake to avoid exceeding the ADI. Paper presented at 'Food Additives—The Changing Climate?', First International Congress, Vienna, 5–7 February.

HALLAS-MØLLER, T., 1995, Using the budget method as a quick screening method identifying food additives for which further monitoring is not warranted on health grounds. Draft report prepared for the SCF, 1 June.

HANSEN, S. C., 1966, Acceptable daily intake of food additives and ceiling on levels of use. *Food and Cosmetics Toxicology*, **4**, 427–432.

HANSEN, S. C., 1979, Conditions for use of food additives based on a budget for an acceptable daily intake. *Journal of Food Protection*, **42**, 429–432.

HARALDSDOTTIR, J., 1993, Minimising error in the field: quality control in dietary surveys. *European Journal of Clinicial Nutrition*, **47** (suppl S2), 19-24.

NATIONAL FOOD AGENCY OF DENMARK, 1986, *Danskernes Kostvaner 1985, 1*. Hovedresultater. Publikation nr. 136, November.

NUTRISCAN, 1994, Options for the routine collection of data on usage levels of food additives in the European Union. Final Report, December.

SCHOELLER, D. A., 1990, How accurate is self-reported dietary intake? *Nutrition Reviews*, **48**, 373–379.

An Evaluation of the Budget Method for Screening Food Additive Intake

ILSI

International Life Sciences Institute

Summary Report

Prepared under the responsibility of ILSI Europe Food Chemical Intake Task Force

AN EVALUATION OF THE BUDGET METHOD FOR SCREENING FOOD ADDITIVE INTAKE

SUMMARY REPORT PREPARED UNDER THE RESPONSIBILITY OF

ILSI EUROPE FOOD CHEMICAL INTAKE TASK FORCE

APRIL 1997

CONTENTS

1. INTRODUCTION

The Budget Method is intended to provide a basis for a simple, inexpensive screening method for additive intake. It relies on assumptions regarding physiological requirements for energy and liquid and on energy density of food, rather than detailed information from food consumption surveys. The theoretical maximum daily intake (TMDI) is calculated by assuming that all foods contributing to energy intake, and all beverages contributing to liquid intake, will contain the additive at maximum permitted use levels. An additive is said to "pass" the Budget Method screen if the calculated TMDI is lower than the additive's corresponding acceptable daily intake (ADI).

Figure 1. Assumptions made in screening additive intake using the Budget Method: Additives used in food.

Maximum energy intake by young children consuming table foods is 100 kcal / kg body weight / day. Maximum energy intake over the course of a lifetime is 50 kcal / kg body weight / day. 50 kcal = 25 g food	25 g food = 1 / 40 kg food
The Theoretical Maximum Daily Intake (TMDI) of an additive can be calculated as: $\frac{\text{Maximum use level (mg additive / kg food)}}{40}$ $\frac{\text{Maximum use level (mg additive / kg food)}}{80}$ $\frac{\text{Maximum use level (mg additive / kg food)}}{160}$	*If:* The additive is used in foods other than baby foods and the maximum amount of food required to meet energy requirements is: $\frac{1/40 \text{ kg}}{\text{kg body weight}}$ 50% of foods consumed are processed foods. The additive is used in processed foods other than those considered to be important in the daily diet (e.g., dairy, meat, fish, poultry, vegetable or cereal products).

An adaptation of the Budget Method developed by Søren Hansen (1966) of the National Food Agency of Denmark has been suggested as an appropriate screening method for determining priorities for monitoring the consumption and use of additives as required under European Union (EU) Directives 94/35/EC, 94/36/EC and 95/2/EC. Before it is accepted in the EU for such a purpose, however, the underlying assumptions should be examined to evaluate the potential for error in the use of results. Clearly, the Budget Method must minimise Type II (false negative) errors which could result in unchecked use of a substance that should have been a priority for monitoring. At the same time, Type I (false positive) errors should be as low as possible to prevent unnecessary expenditure of time and resources in pursuit of more detailed intake estimates.

This report summarises work performed by TAS (Technical Assessment Systems) International at the request of the International Life Sciences Institute - ILSI Europe Food Chemical Intake Task Force. The aim of the report was to evaluate the validity of assumptions on which the Budget Method is based and to assess the potential for Type I and II errors in using the method to establish additive monitoring priorities. The report should be read in conjunction with the references cited herein and listed at the end of this report.

Figure 2. Assumptions made in screening additive intake using the Budget Method: Additives used in beverages.

<table>
<tr><td colspan="2">Maximum intake of liquids (other than milk) is 100 ml/kg body weight/day.

100 ml = 1/10 litre</td></tr>
<tr><td>The Theoretical Maximum Daily Intake (TMDI) of an additive can be calculated as:

$\frac{\text{Maximum use level (mg additive/l beverage)}}{10}$

$\frac{\text{Maximum use level (mg additive/l beverage)}}{40}$</td><td>If:

The additive is used in all non-milk beverages and the maximum amount of non-milk beverages consumed is:
$\frac{\text{1/10 litre}}{\text{kg body weight}}$

The additive is used only in soft drinks and maximum soft drink intake is 25% of non-milk beverage intake.</td></tr>
</table>

2. BUDGET METHOD

The Budget Method was designed to compare food additive ADIs with "ceilings of use" calculated on the basis of maximum consumption of food and beverages potentially containing the additives (Hansen 1966, 1979). In budget calculations for additives used in both solid foods and beverages, the ADIs are split into two fractions. The proportion of the ADI allocated to food and the proportion allocated to beverages are decided arbitrarily to accommodate technological requirements.

The proposal for use of the Budget Method calls for calculation of TMDIs using the assumptions outlined in Figures 1 and 2. Figure 1 describes the factors used and the assumptions made in estimating intake of additives from solid foods. Figure 2 describes the factors used and the assumptions made when estimating intake of additives from beverages.

3. METHODS

TAS International examined data from nationwide food consumption surveys conducted in the United Kingdom, former West Germany, and the United States (Figure 3). It examined whether Budget Method assumptions regarding energy intake, energy density of food and beverage soft drink consumption provide a valid basis for screening additive intake. Because the Budget Method assumes that ADIs for additives relate to average intake over a lifetime, the validity of assumptions was assessed on the basis of population averages rather than on high-level intake.

Because of limitations posed by the availability of European survey data, Budget Method assumptions regarding proportions of the diet accounted for by processed foods potentially containing additives could not be tested. The UK surveys were conducted on specific age groups, and results therefore were summarised on the basis of the UK age groups and on broad food groups.

4. CASE STUDIES

To evaluate the potential for Type I and II errors in using the Budget Method, case studies were conducted for two hypothetical food additives with different characteristics and proposed uses (Tables 1 and 2).

Table 1. Case Study of Additive 1: Used at similar levels in a broad range of foods consumed by a significant portion of the population.

Food Category	*Use limit (mg/kg food)*
Breads	50
Baked Products	50
Pastas	30
Cereals	50
Rice	30

Table 2. Case Study of Additive 2: Used in varying concentrations in a range of foods consumed by specific segments of the population.

Food Category	*Use limit (mg/kg food or mg/l beverage)*
Soft Drinks	350
Biscuits	1000
Confectionery	500

Maximum permitted use levels in specific food groups were selected to reflect typical use levels.

Food Additive 1 was designed to be representative of an additive used at similar levels in a broad range of foods

Figure 3. Food consumption survey data used in Budget Method validation.

United Kingdom

Data used in the analyses were taken from the following summary sources of survey data published by the Ministry of Agriculture, Fisheries and Food:

- *Food and Nutrient Intakes of British Infants Ages 6–12 Months;* 1986; 488 infants; based on 7-day food intake records.
- *National Diet and Nutrition Survey: Children Ages 1.5–4.5 years;* July 1992 and June 1993; 1,675 children; based on 4-day food intake records.
- *The Diets of British School Children;* 1983; 3,581 children ages 10–11 and 14–15; based on 7-day food intake records.
- *The Dietary and Nutritional Survey of British Adults;* October 1986 - August 1987; 2,197 adults ages 16–64; based on 7-day food intake records.

Former West Germany

National Consumption Study (NVS); October 1985–January 1989. Seven-day weighed-intake data were collected from more than 25,000 individuals 4 years of age and older, using a system consisting of more than 6,000 food codes. Data used in the present analyses were taken from a dataset containing records of average daily intake of foods in 90 summary food groups by individual survey respondents.

United States

US Department of Agriculture 1989–90, 1990–91 and 1991–92 Continuing Surveys of Food Intake by Individuals (CSFII). Together, the CSFII surveys measured dietary intake of more than 11,000 individuals over a 3-day period. Although these data clearly do not directly reflect European food intake patterns, it was believed that the extensive information on intake by individuals could be of potential value in examining basic Budget Method assumptions regarding food additive intake. For example, neither the UK surveys nor the German survey collected data on individuals in all age groups; the US data provide supplementary data on total population intake.

consumed by a significant proportion of the population. Food Additive 2 was designed to be representative of an additive used in varying concentrations in a range of foods consumed by specific segments of the population (e.g., an intense sweetener).

The Budget Method TMDI for Additive 2 was calculated allocating half of the budget to soft drinks and half to "high additive" solid foods (biscuits and confectionery), using a liquid intake factor of 40 and a food consumption factor of 160 (Figures 1 and 2).

Food consumption survey–based intake estimates for Additives 1 and 2 were calculated as precisely as possible given the limitations of the survey data. Additive use levels were applied to food consumption data in the categories shown in Tables 1 and 2.

Limitations on the availability of the survey data used in these analyses precluded calculation of per-user intake. However, it was assumed that all people will be consumers of the additives in question at some point in their lifetimes, and *per capita* estimates were therefore judged to be representative of lifetime intake. Because it is generally believed that the Budget Method generates conservative estimates of intake, TMDIs were compared with 95th percentile *per capita* intake. Where survey data did not permit assessment of 95th percentile *per capita* intake, the intake values were estimated using the Bernier method (Bernier et al. 1994) at three times the mean.

5. RESULTS

5.1 Energy intake

The survey results indicated that the Budget Method assumptions of 100 kcal/kg body weight energy intake for 1-year-olds and of 50 kcal/kg body weight for the general population overestimate actual lifetime average energy intake by a small margin.

5.2 Energy density of food

In developing the Budget Method, Hansen (1966) assumed that 50 g of food have an energy value equal to 100 kcal, which is equivalent to 2 kcal/g. In contrast, UK and US survey data indicate that the mix of food consumed, on average, is of slightly lower energy density (the German data did not permit calculation of energy density).

The Budget Method is therefore based on an overestimate of energy density of foods consumed by the general population. The extent of the overestimate would be even greater for subpopulations consuming large quantities of low-calorie foods.

5.3 Beverage and soft drink consumption

Budget Method calculations for additive intake from beverages are based on the assumption that 100 ml/kg body weight of liquids, excluding milk, represents maximum consumption. The maximum soft drink consumption is assumed to be 25% of this consumption, or 25 ml/kg body weight/day.

Beverage consumption data are shown for selected age groups in Table 3. In most cases these are clearly overestimated by the Budget Method assumptions. It is likely that UK soft drink consumption by 1.5–4.5-year-old children is higher than that by children of comparable ages in Germany and the USA because water-diluted rose hip, blackcurrant and other fruit cordials popular with young children in the UK are included in summarised soft drink consumption data.

5.4 Comparison of TMDIs with Additive Intake Estimates Based on Food Consumption Survey Data

Budget Method TMDIs calculated for Food Additives 1 and 2 (0.63 and 15.00 mg/kg body weight) are above the 95th percentile *per capita* estimates of intake for all three national surveys investigated (Table 4), indicating that the budget method provides a conservative first estimate of intake.

It is unlikely that Type I or II errors would occur in the assignment of monitoring priority for additives like Food Additive 1 using the Budget Method. Additive 1 was intended to be representative of an additive used at similar levels in a broad range of foods consumed by a significant proportion of the population.

The scenario for additives like Food Additive 2 is different. This hypothetical additive was intended to be representative of an additive used in varying concentrations in a range of foods consumed by specific segments of the population (e.g., an intense sweetener). The Budget Method TMDI for Food Additive 2 is three to five times the survey-based estimates of intake, indicating a potential for Type I (false positive) errors. This means that the Budget Method is likely to falsely identify this type of additive as warranting further attention.

Table 3. Beverage and soft drink consumption estimates based on UK, West German and US survey data.

Age Group	*Mean Per Capita Intake of Non-Milk Beverages and of Soft Drinks (g/kg body weight/day)*					
	UK		West German		US	
	All non-milk beverages	Soft drinks only	All non-milk beverages	Soft drinks only	All non-milk beverages	Soft drinks only
Total population	NA[1]	NA	17[2]	2[2]	13	4
6 - 12 months	18	4	NA	NA	17	2
1.5–4.5 years	31	25[3]	NA	NA	18[4]	5[4]
10 –11 years	8	4	16	5	11	5
14–15 years	9	3	13	4	11	6
16–64 years	23	2	18	2	13	4

(1) NA = not applicable (population group not included in survey).
(2) Ages 4+.
(3) Water-diluted rose hip, blackcurrant, and other fruit cordials popular with young children in the UK are included in summarised soft drink intake data.
(4) Ages 1–5 years.

Table 4. Budget Method TMDI and survey-based intake estimates for case study Additives 1 and 2.

Food Additive	*Budget Method TMDI (mg/kg body weight)*	*Intake Estimates Based on Food Consumption Survey Data*	
		Population	95th percentile *per capita* intake (mg/kg body weight/day)
1	0.63	UK[1]	0.34
		West German[2]	0.40
		US	0.43
2	15.00	UK	2.88
		West German	4.42
		US	5.12

(1) Ages 16–64; 95th percentile intake for this population computed as (mean intake x 3), as recommended by Bernier et al. (1994).
(2) Ages 4+.

6. CONCLUSIONS

The Budget Method assumptions regarding energy intake, energy density of foods, and beverage and soft drink consumption using UK, German and US national food consumption data were reviewed, and the following conclusions were reached:

- Assumptions regarding energy intake and beverage and soft drink consumption by the general population are overestimates of actual average levels. These probably add to the overall conservatism of the Budget Method.

- The Budget Method assumption regarding the energy density of foods may result in small overestimates. This probably detracts slightly from the overall conservatism of the method.

In the two case studies, Budget Method TMDIs were larger than survey-based *per capita* additive intake estimates, providing evidence that the Budget Method produces overestimates of additive intake. The potential for Type II (false negative) errors is therefore judged to be small whereas the potential for Type I (false positive) errors seems to be relatively large.

Thus, the Budget Method appears to be a satisfactory conservative first screen for establishing priorities for monitoring the intake of food additives, based on potential average lifetime food consumption.

7. REFERENCES

Bernier JJ, Causeret J, Hoellinger H, Suschetet M. Méthode pour assurer, au sein d'une population, le non-dépassement de la dose journalière admissible des additifs alimentaires. *Med Nutri* 1994;1:9–17

Douglass JS, Barraj LM, Tennant DR, et al. Evaluation of the budget method for screening food additive intakes. *Food Addit Contam* 1997 (in press)

Hallas-Møller T. Using the budget method as a quick screening method for identifying food additives for which further monitoring is not warranted on health grounds. Draft report prepared for the SCF, 1 June 1995

Hansen SC. Acceptable daily intake of food additives and ceiling on levels of use. *Food Cosmet Toxicol* 1966;4:427–432

Hansen SC. Conditions for use of food additives based on a budget for an acceptable daily intake. *Food Protect* 1979;42:429–432

The International Life Sciences Institute (ILSI) is a nonprofit, worldwide foundation established in 1978 to advance the understanding of scientific issues relating to nutrition, food safety, toxicology, and the environment.

By bringing together scientists from academia, government, industry and the public sector, ILSI seeks a balanced approach to solving problems of common concern for the well-being of the general public. ILSI is affiliated with the World Health Organization as a nongovernmental organisation and has specialized consultative status with the Food and Agricultural Organization of the United Nations.

Headquartered in Washington, D.C. USA, ILSI has branches in Argentina, Australasia, Brazil, Europe, India, Japan, Korea, Mexico, North America, Southeast Asia, Thailand and a focal point in China.

ILSI Europe
Avenue E. Mounier, 83, Box 6
B-1200 Brussels
BELGIUM
Telephone: (+32) 2 771 0014
Telefax: (+32) 2 762 0044

List of Appendices

1) General Requirements
2) Product information
3) Methods of analysis used to determine compliance with the specifications
4) Documentation regarding the manufacturing process
5) Safety documentation
6) Documentation regarding the production microorganism

Appendix 1

General Requirements

1. Formal request for treatment of confidential commercial information (CCI)
2. Checklist for General requirements
3. Checklist for applications for substances added to food

Appendix 1.1

Formal request for treatment of confidential commercial information (CCI)

Novozymes respectfully request that the selected and marked parts of Appendix 6 are treated as confidential commercial information (CCI).

The documents in Appendix 6 contain detailed description of the construction of the genetically modified production strain and the introduced DNA. While individual steps in the DNA construction might be well known or publicly available information, the exact steps and sequence of those constitutes information that represent the state-of-the-art of one of Novozymes' core technologies, which has been obtained as a result of substantial investment in research and development within rDNA technology. Therefore, the selected parts of Appendix 6 are claimed confidential for an unlimited period of time.

June 2018

Alexander Striebeck
Regulatory Specialist
Regulatory Affairs
Novozymes A/S

Appendix 1.2

Checklist for General requirements

General requirements (3.1.1)		
Check	**Page No.**	**Mandatory requirements**
☒	4	A Form of application ☒ *Application in English* ☒ *Executive Summary (separated from main application electronically)* ☒ *Relevant sections of Part 3 clearly identified* ☒ *Pages sequentially numbered* ☒ *Electronic copy (searchable)* ☒ *All references provided*
☒	8	B Applicant details
☒	9	C Purpose of the application
☒	9	D Justification for the application ☒ *Regulatory impact information* ☒ *Impact on international trade*
☒	10	E Information to support the application ☒ *Data requirements*
☒	11	F Assessment procedure ☒ *General* ☐ *Major* ☐ *Minor* ☐ *High level health claim variation*
☒	11 11 27	G Confidential commercial information ☒ *CCI material separated from other application material* ☒ *Formal request including reasons* ☒ *Non-confidential summary provided*
☒	11	H Other confidential information ☐ *Confidential material separated from other application material* ☐ *Formal request including reasons*
☒	11	I Exclusive Capturable Commercial Benefit ☐ *Justification provided*
☒	11	J International and other national standards ☒ *International standards* ☒ *Other national standards*
☒	12	K Statutory Declaration
☒		L Checklist/s provided with application ☒ *3.1.1 Checklist* ☒ *All page number references from application included* ☒ *Any other relevant checklists for Chapters 3.2–3.7*

Appendix 1.3

Checklist for applications for substances added to food

Processing aids (3.3.2)		
Check	**Page No.**	**Mandatory requirements**
☒	13	A.1 Type of processing aid
☒	14	A.2 Identification information
☒	15	A.3 Chemical and physical properties
☒	15	A.4 Manufacturing process
☒	19	A.5 Specification information
☒	20	A.6 Analytical method for detection
☐		B.1 Industrial use information (chemical only)
☐		B.2 Information on use in other countries (chemical only)
☐		B.3 Toxicokinetics and metabolism information (chemical only)
☐		B.4 Toxicity information (chemical only)
☐		B.5 Safety assessments from international agencies (chemical only)
☒	20	C.1 Information on enzyme use on other countries (enzyme only)
☒	21	C.2 Toxicity information of enzyme (enzyme only)
☒	23	C.3. Allergenicity information of enzyme (enzyme only)
☒	24	C.4. Overseas safety Assessment Reports
☒	25	D.1 Information on source organism (enzyme from microorganism only)
☒	25	D.2 Pathogenicity and toxicity of source microorganism (enzyme from microorganism only)
☒	26	D.3 Genetic stability of source organism (enzyme from microorganism only)
☒	26	E.1 Nature of genetic modification of source organism (enzyme from GM source microorganism)
☒	28	F.1 List of foods likely to contain the processing aid
☒	28	F.2 Anticipated residue levels in foods
☒	30	F.3 Information on likely level of consumption
☒	30	F.4 Percentage of food group to use processing aid
☒	31	F.5 Information on residues in foods in other countries (if available)
☒	31	F.6 Where consumption has changed, information on likely consumption

Appendix 2

Product information

1. Attenuzyme Core application sheet
2. Product Data Sheet for Attenuzyme Core
3. DK approval certificate for baking applications (GoldCrust 3300 BG)
4. DK approval certificate for syrup, brewing, and distilling applications (Spirizyme Flex)

Attenuation control and light beer production

Novozymes offers a broad range of attenuation control products to allow brewers to manage fluctuations in attenuation and create highly attenuated beers in a simple, cost-effective manner.

Benefits

- Customized solutions to reach specific attenuation targets and needs
- Manage consumer trends towards low-calorie and light beers
- Opportunity to reach desired maltose to glucose ratios
- Obtain the desired level of fermentable sugars every time

Products

Five products can be used—either alone or combined—to achieve desired attenuation levels within the specified production parameters of the individual brewery and brand:

Attenuzyme® Pro is a high-performing blend of glucoamylase and pullulanase that makes it possible to hit high attenuation targets in short reaction times, taking advantage of the synergy between these two enzyme activities during the hydrolysis of amylopectin and amylose.

Attenuzyme® Core is a glucoamylase for producing highly fermentable glucose-based worts.

AMG® 300 L BrewQ is a classic glucoamylase for producing highly fermentable glucose-based worts.

Novozym® 26062 is a pullulanase that accelerates attenuation and can be applied for a moderate increase in the attenuation of maltose-based wort.

Fungamyl® BrewQ is a maltogenic alpha-amylase used to breakdown of starches, facilitating a higher alcohol output.

Performance

The Novozymes range of attenuation enzymes help to break down amylose and amylopectin.

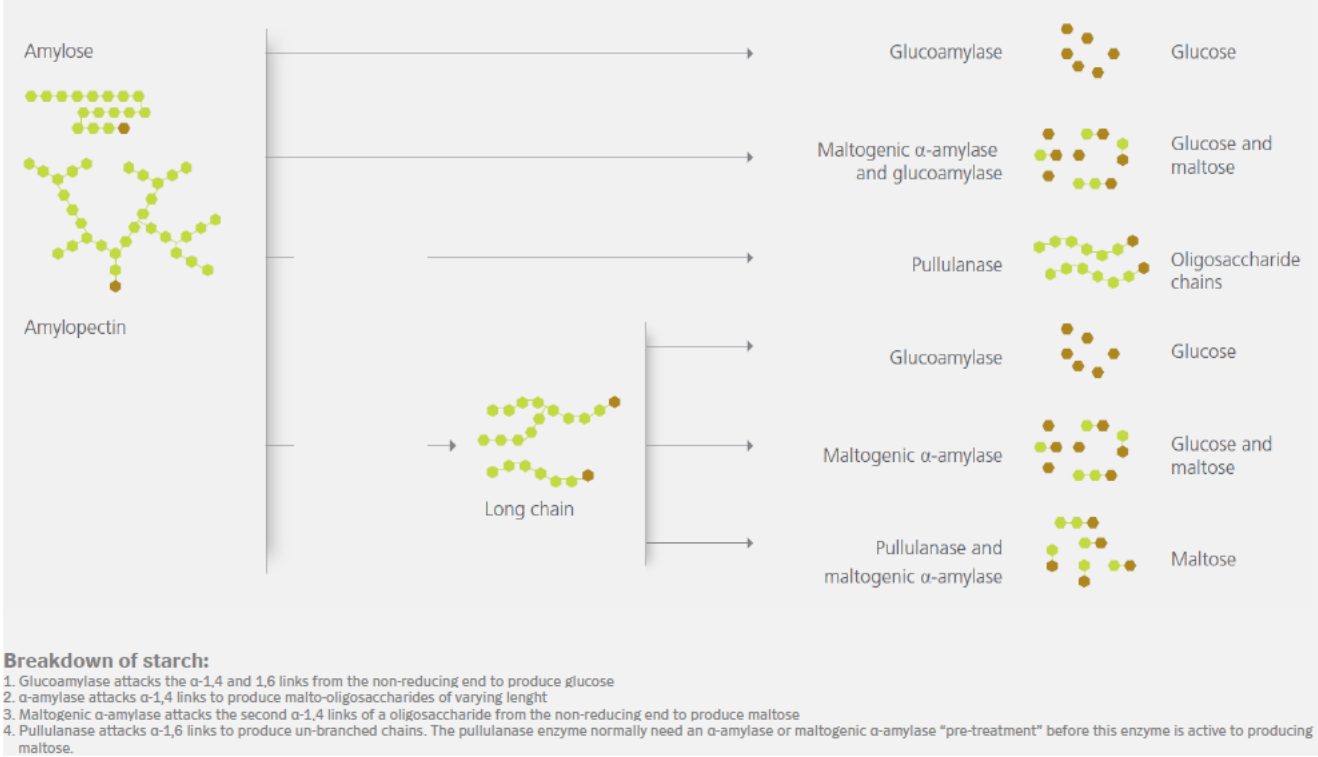


Fig. 1. Amylopectin breakdown to glucose and maltose by glucoamylase, α-amylase and pullulanase

The desired attenuation solution is a choice for the brewery to make. This application sheet focuses on the performance of Attenuzyme® Pro and Attenuzyme® Core, which are the recommended solutions for highly attenuated beer.

Rates of production of fermentable sugars are dose- and time-dependent.

Formed fermentable sugars [%] as function of the saccharification time

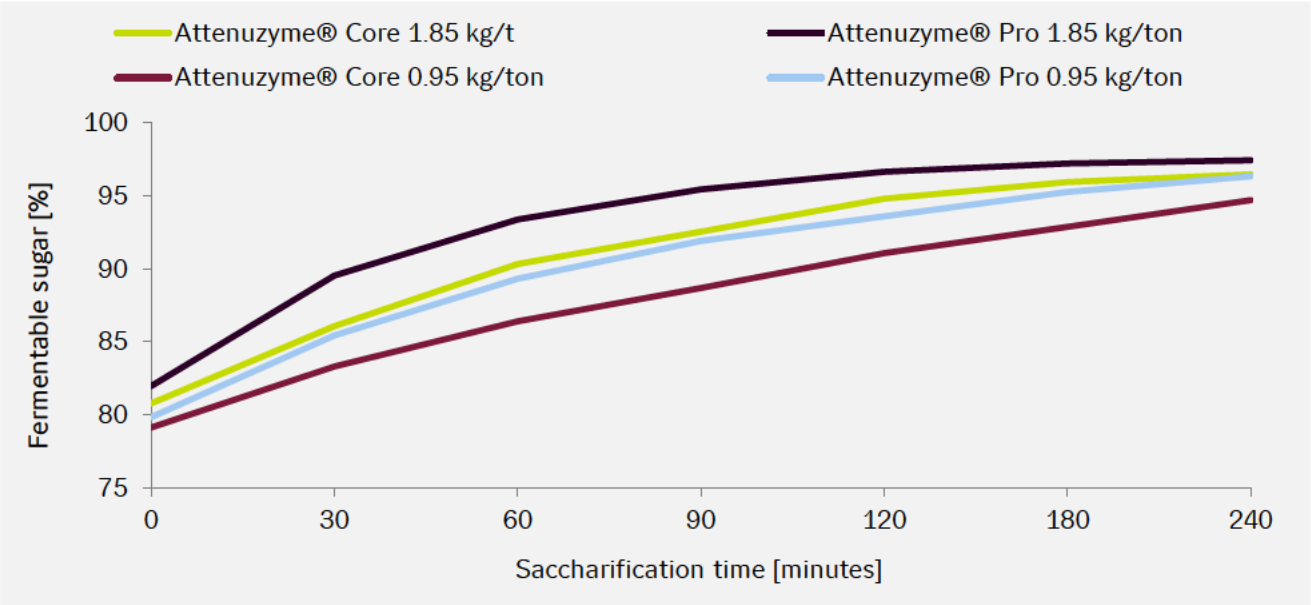


Fig. 2. Time response curves for Attenuzyme® Core and Attenuzyme® Pro on a 50% malt and 50 % corn grist composition at 64ºC, pH 5.4

As a rule of thumb, the RDF (real degree of fermentation) correlates with the amount of fermentable sugars * 0.91.
As shown in figure 2, Attenuzyme® Pro is faster at degrading dextrins into fermentable sugars when compared to Attenuzyme® Core due to additional pullulanase activity in Attenuzyme® Pro. The pullulanase works in synergy with malt β-amylase, generating a higher maltose content than Attenuzyme® Core and shortening saccharification times.

Usage

Dosage response curve for Attenuzyme® Core

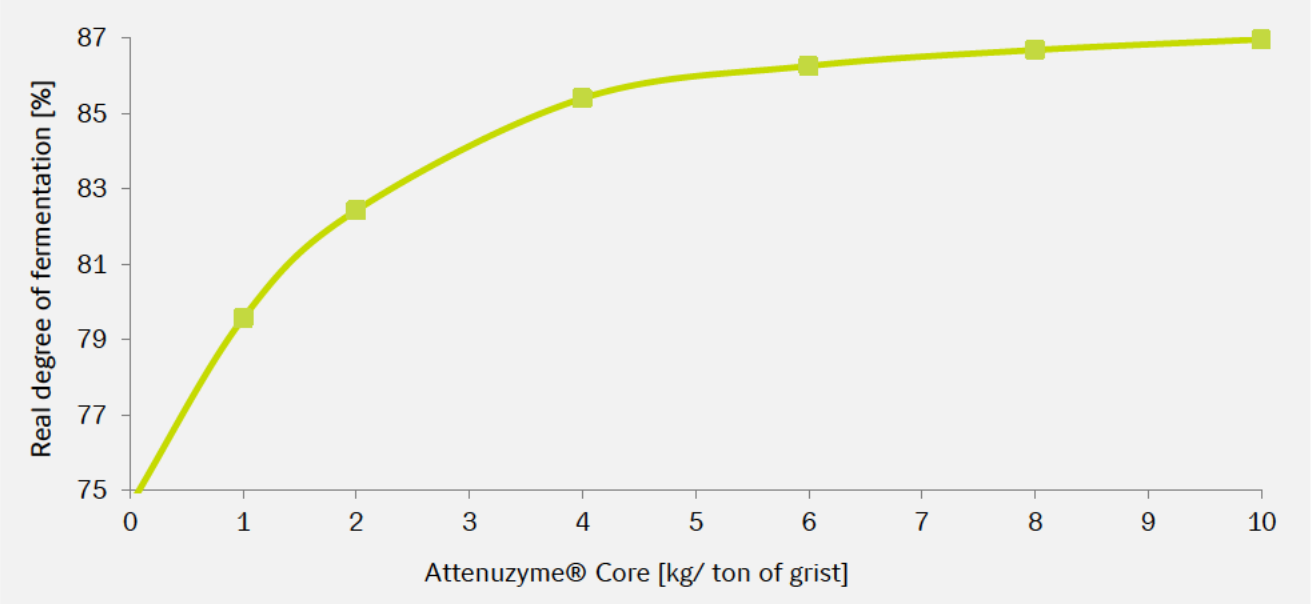


Fig. 3. Dosage response (real degree of fermentation [%]) of Attenuzyme® Core at 64°C after 60 minutes

Dosage response curve for Attenuzyme® Pro

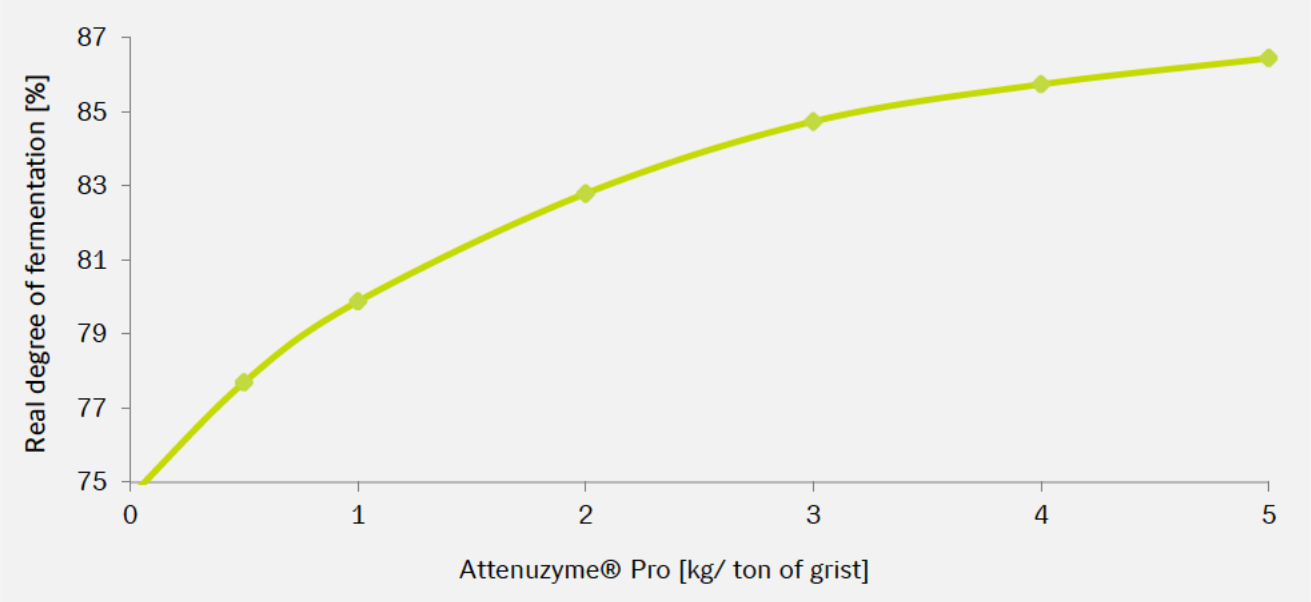


Fig. 4. Dosage response (real degree of fermentation [%]) of Attenuzyme® Pro at 64°C after 60 minutes

Application examples and dosage suggestions

Desired attenuation [%]		Option	Enzymes	Dosage range	Units (per ton grist or hL beer)	Point of addition
RDF	ADF					
70–75	85–90	A	Fungamyl® BrewQ	0.5 to 5	g/hL	Start of fermentation
75–80	90–95	A	AMG® 300 L BrewQ	1.2 to 3.5	kg/ton	Mashing-in
			+ Novozym® 26062	2.4 to 3.6	kg/ton	
		B	Attenuzyme® Core	0.35 to 1	kg/ton	Mashing-in
		C	Attenuzyme® Core	0.25 to 0.75	kg/ton	Mashing-in
			+ Novozym® 26062	1.2 to 2.4	kg/ton	
		D	Attenuzyme® Pro	0.3 to 1	kg/ton	Mashing-in
80–90	95–100	A	Fungamyl® BrewQ	4 to 8	g/hL	Start of fermentation
			+ Novozym® 26062	1.2 to 3.6	kg/ton	Mashing-in
		B	Fungamyl® BrewQ	2.4 to 4.8	g/hL	Start of fermentation
			+ Novozym® 26062	12 to 18	kg/ton	
		C	AMG® 300 L BrewQ	6 to 18	kg/ton	Mashing-in or hot wort (63°C)
			+ Novozym® 26062	6 to 18	kg/ton	
		D	Attenuzyme® Core	2 to 6	kg/ton	Mashing-in or hot wort (63°C)
		E	Attenuzyme® Core	1.5 to 5	kg/ton	Mashing-in or hot wort (63°C)
			+ Novozym® 26062	2.4 to 4.8	kg/ton	
		F	Attenuzyme® Pro	0.25 to 5	kg/ton	Mashing-in or hot wort (63°C)

Table 1. How to adjust fermentability

Attenuzyme® products are inactivated during wort boiling. In case of use during fermentation, normal pasteurization temperatures will not inactivate the enzymes, but this may cause post-sweetening of the beer.

Please contact our global Technical Service team for further suggestions on dosage during fermentation.

More information can be found in the “Brewing Handbook” and “Brewing Calculator,” both of which are available at Novozymes Market.

Optimum pH and temperature

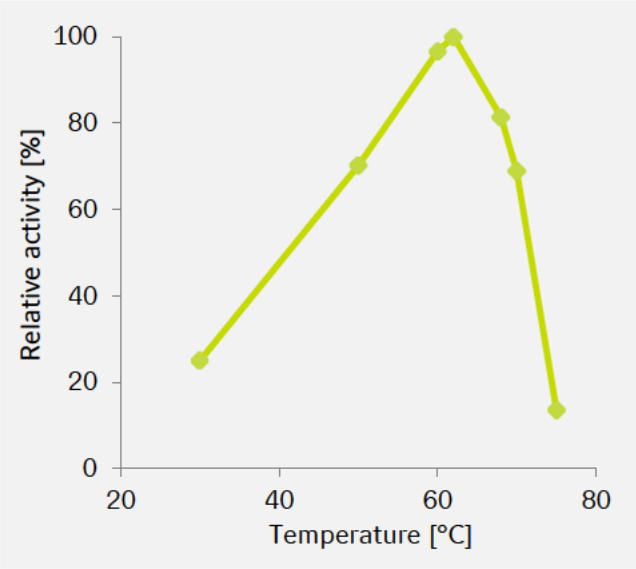


Fig. 5. Temperature activity profile of Attenuzyme®

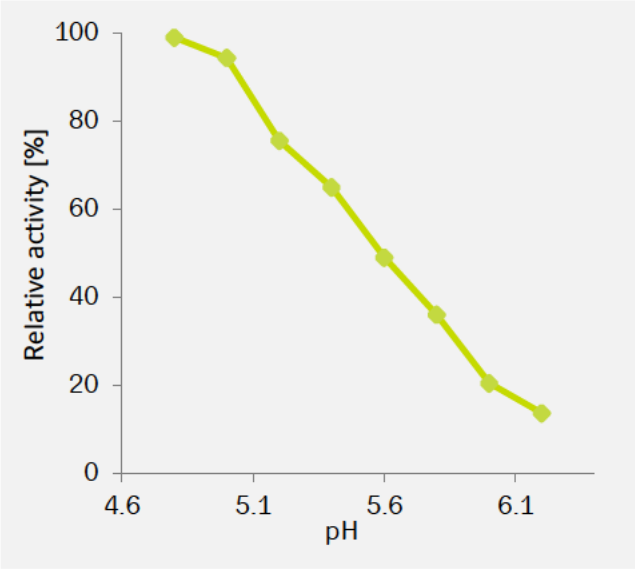


Fig. 6. pH activity profile of Attenuzyme®

Product data

Attenuzyme® Pro	
Declared enzyme	A multi-component enzyme solution comprised of a fungal alpha-amylase, glucoamylase, and a patented pullulanase technology for accelerated production of highly fermentable glucose-based worts
Catalyzes the following reactions:	Glucoamylase that hydrolyzes (1,4)- and (1,6)-alpha-D-glucosidic linkages at the non-reducing ends of polysaccharides to produce glucose. Pullulanase that hydrolyzes (1,6)-alpha-D-glucosidic linkages in pullulan, amylopectin and glycogen to produce smaller fragments of linear dextrin
Declared activity	1300 AGU/g & 315 PUN/g
E.C./I.U.B.no.:	3.2.1.3 & 3.2.1.41
Physical form	Liquid
Production method	The enzyme product is manufactured via fermentation of a microorganism not present in the final product. The production organism is improved by means of modern biotechnology
Density	1.15 g/ml

Attenuzyme® Core	
Declared enzyme	Glucoamylase (glucan 1,4-alpha-glucosidase)
Catalyzes the following reactions:	Hydrolyzes (1,4)- and (1,6)-alpha-D-glucosidic linkages at the non-reducing ends of polysaccharides to produce glucose
Declared activity	1600 AGU/g
E.C./I.U.B.no.:	3.2.1.3
Physical form	Liquid
Production method	The enzyme product is manufactured via fermentation of a microorganism not present in the final product. The production organism is improved by means of modern biotechnology
Density	1.13 g/ml

AMG® 300 L BrewQ	
Description	A classic glucoamylase (amyloglucosidase), used for production of highly fermentable, glucose-based worts
Declared enzyme	Glucoamylase (glucan 1,4-alpha-glucosidase)
Catalyzes the following reactions:	Hydrolyzes (1,4)- and (1,6)-alpha-D-glucosidic linkages at the non-reducing ends of polysaccharides to produce glucose
Declared activity	300 AGU/ml
E.C./I.U.B.no.:	3.2.1.3
Physical form	Liquid
Production method	The enzyme product is manufactured via fermentation of a microorganism not present in the final product. The production organism is not modified using modern biotechnology
Density	1.17 g/ml

Novozym® 26062	
Description	A heat-stable pullulanase that accelerates production of highly fermentable worts when used in conjunction with a glucoamylase
Declared enzymes	Pullulanase
Catalyzes the following reactions:	Hydrolyzes (1,6)-alpha-D-glucosidic linkages in pullulan, partially degraded amylopectin and partially degraded glycogen to produce smaller fragments of linear dextrin
Declared activities	400 PUN/g
E.C./I.U.B.no.:	3.2.1.41
Physical form	Liquid
Production method	The enzyme product is manufactured via fermentation of a microorganism not present in the final product. The production organism is improved by means of modern biotechnology
Density	1.20 g/ml

Fungamyl® BrewQ	
Description	A classic fungal alpha-amylase used to increase the breakdown of starches, facilitating higher alcohol output
Declared enzymes	Alpha-amylase
Catalyzes the following reactions:	Endo-amylase that hydrolyzes (1,4)-alpha-D-glucosidic linkages in starch polysaccharides
Declared activities	800 FAU-F/g
E.C./I.U.B.no.:	3.2.1.1
Physical form	Liquid
Production method	The enzyme product is manufactured via fermentation of a microorganism not present in the final product. The production organism is improved by means of modern biotechnology
Density	1.26 g/ml

More information about the products is available at Novozymes Market.

Stability

Please see the Product Data Sheet at Novozymes Market.

Safety, handling and storage

Safety, handling and storage guidelines are provided with all products.

Get ahead

Staying ahead of the dynamic food and beverage market requires the best technology and expertise to become even more flexible, efficient and profitable. With our solutions and expertise, Novozymes can support you on that journey. Let's transform the quality and sustainability of your business together.

About Novozymes
Novozymes is the world leader in biological solutions. Together with customers, partners and the global community, we improve industrial performance while preserving the planet's resources and helping build better lives. As the world's largest provider of enzyme and microbial technologies, our bioinnovation enables higher agricultural yields, low-temperature washing, energy-efficient production, renewable fuel and many other benefits that we rely on today and in the future. We call it Rethink Tomorrow.

Novozymes Switzerland AG
Neumattweg 16
4243 Dittingen
Switzerland
foodandbeverages@novozymes.com

Novozymes A/S
Krogshoejvej 36
2880 Bagsvaerd
Denmark

Jan, 2017 · Luna No. 2013-07877-02

Attenuzyme® Core

In this product the key enzyme activity is provided by glucoamylase that hydrolyzes (1,4)- and (1,6)-alpha-D-glucosidic linkages at the non-reducing ends of polysaccharides

PRODUCT CHARACTERISTICS/PROPERTIES

Component name	Glucoamylase (glucan 1,4-alpha-glucosidase)
Activity	1600 AGU/g
Side activities	The product contains controlled activity of Alpha-amylase
Color	Light to dark brown
Physical form	Liquid
Approximate density (g/ml)	1.13

Color can vary from batch to batch. Color intensity is not an indication of enzyme activity.

PRODUCT SPECIFICATION

	Lower Limit	Upper Limit	Unit
Amyloglucosidase unit AGU	1600		/g
pH at 25°C	3	5	
Total viable count	-	10000	/g
Molds	-	100	/g
Yeast	-	100	/g
Coliform bacteria	-	30	/g
E.coli	Not Detected		/25 g
Beer spoilers bacteria	-	100	/g
Salmonella	Not Detected		/25 g
Heavy metals		Max 30	mg/kg
Lead		Max 5	mg/kg
Arsenic		Max 3	mg/kg
Cadmium		Max 0.5	mg/kg
Mercury		Max 0.5	mg/kg

The enzyme analytical method is available from the Customer Center or sales representative.

COMPOSITION

Preservatives	Potassium sorbate Sodium benzoate
Stabilizers	Sucrose/Glucose, D-

ALLERGEN

Allergen	Substance contained[1]	Allergen	Substance contained[1]
Celery	no	**Molluscs**	no
Cereals containing gluten[2/4]	no	**Mustard**	no
Crustaceans	no	**Nuts[3]**	no
Egg	no	**Peanuts**	no
Fish	no	**Sesame**	no
Lupin	no	**Soy**	no
Milk (including lactose)	no	**Sulphur dioxide/sulphites, more than 10 mg per kg or l**	no

[1]Definition of substances according to EU Regulation 1169/2011, as amended. List covers allergens mentioned in 21 USC 301 (US) and GB 7718-2011 (China).
[2]i.e.wheat, rye, barley, oats, spelt, kamut
[3]i.e. almond, hazelnut, walnut, cashew, pecan nut, Brazil nut, pistacchio nut, macadamia nut and Queensland nut
[4] If No: Glutenfree i,e. < 20ppm (EU Regulation 828/2014)

NUTRITIONAL VALUES

The product has a typical nutritional value of approximately 735 kJ/100 g enzyme product.

• **Protein**		42 g/100 g
• **Carbohydrate**		1 g/100 g
• **Organic acid**		0 g/100 g
• **Ash**		0 g/100 g
- Sodium	(0.05 g/100 g)	
• **Moisture**		57 g/100 g

GM STATUS

This product is not a GMO.

Production organism Aspergillus niger

The enzyme product is manufactured by fermentation of a microorganism that is not present in the final product. The production organism is improved by means of modern biotechnology.

STORAGE CONDITION

Recommended storage: 0-10 °C (32-50 °F)

Packaging must be kept intact, dry, and away from sunlight. Please follow the recommendations and use the product before the best before date to avoid the need for a higher dosage.

Best before: You will find the best before date in the certificate of analysis or on the product label.

The product gives optimal performance when stored as recommended and used prior to the best-before date.

The product can be transported at ambient temperature. Following delivery, the product should be stored as recommended.

SAFETY AND HANDLING PRECAUTIONS

Enzymes are proteins. Inhalation of dust or aerosols may induce sensitization and may cause allergic reactions in sensitized individuals. Some enzymes may irritate the skin, eyes, and mucous membranes upon prolonged contact. See the MSDS or Safety Manual for further information regarding safe handling of the product and spills.

COMPLIANCE

The product complies with the recommended purity specifications for food-grade enzymes given by the Joint FAO/WHO Expert Committee on Food Additives (JECFA) and the Food Chemical Codex (FCC).

Kosher certificate is available from the Customer Center or sales representative.

CERTIFICATIONS

Novozymes is a signatory to United Nations Global Compact, United Nations Convention on Biological Diversity and report on our sustainability performance through Global Reporting Initia ive (GRI). See all our commitments under sustainability on www.novozymes.com.

FOOD SAFETY

Novozymes has carried out a hazard analysis and prepared an HACCP plan describing the critical control points (CCPs). The HACCP plan is supported by a comprehensive prerequisite program implemented in Novozymes' GMP practices. The product is produced according to Novozymes' HACCP plan, GMP practices, and additional requirements controlled by Novozymes' Quality Management System.

The product complies with FAO/WHO JECFA- and FCC-recommended purity requirements regarding mycotoxins.

PACKAGING

The product is available in different types of packaging. Please contact the sales representative for more information.

Novozymes A/S
Krogshoejvej 36
2880 Bagsvaerd
Denmark

Ministry of Food, Agriculture and Fisheries
Danish Veterinary and Food Administration

TO WHOM IT MAY CONCERN

DIVISION OF
FOOD QUALITY, TECHNOLOGY
AND MARKETING PRACTICES

20.01.2011

GoldCrust 3300 BG

The Danish Veternary and Food Administration hereby certifies having accepted in 2009 the enzyme product GoldCrust 3300 BG from Novozymes A/S. The product, which is derived from a genetically modified strain of *Aspergillus niger* (BE) expressing the gene encoding glucoamylase from *Talaromyces emersonii*, has been accepted to be used in baking applications, including production of bread and fine bakery at a level of up to 500 AGU per kg flour.

The evaluation of the safety of GoldCrust 3300 BG bas been made in accordance with the principles laid down in the Guidelines for the presentation of data on food enzymes, "cf. Reports of the Scientific Commettee of Food, 27th, Series; EUR 14181, 1992.

Yours faithfully

Birgit Christine Bønsager
Direct tel. +45 33956171
E-mail bicb@fvst.dk

Ministry of Food, Agriculture and Fisheries
Danish Veterinary and Food Administration

TO WHOM IT MAY CONCERN

DIVISION OF
FOOD QUALITY, TECHNOLOGY
AND MARKETING PRACTICES

07.02.2008
File: 2008-20-25-04518/BICB

Spirizyme Flex

The Danish Veternary and Food Administration hereby certified having accepted in 2007 the enzyme product Spirizyme Flex also called Saczyme from Novozymes A/S. The product, which is derived from a genetically modified strain of *Aspergillus niger* expressing the glucoamylase gene from *Talaromyces emersonii*, has been accepted to be used in the production of glucose syrup, alcohol and beer up to a level of 750 AGU per kg starch solid.

The evaluation of the safely of Spirizyme Flex/Saczyme has been made in accordance with the principles laid down in the Guidelines for the presentation of data on food enzymes, "cf. Report of the Scientific Committee of Food, 27th Series, EUR 14181, 1992.

Yours faithfully

Birgit Bønsager

Birgit Christine Bønsager
Ph.D. in Biochemistry
Direct tel. +45 33956171
E-mail bicb@fvst.dk

Appendix 3

Methods of analysis used to determine compliance with the specifications

1. Enzyme activity, AGU/g
2. Heavy metals
3. Total aerobic viable count
4. Total coliforms
5. *Escherichia coli (E. coli)*
6. Salmonella
7. Antimicrobial activity
8. Mycotoxins
9. Absence of production strain

Analysis of Heavy Metals, Lead, Arsenic, Cadmium, and Mercury were performed at an external laboratory, Danish Technological Institute (DTI), Denmark.

AGU, Amyloglucosidase determination

The method is used to measure amyloglucosidase activity in samples from Novozymes.

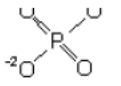
Principle

The analysis principle is described by three reaction steps:

Step 1 is an enzyme reaction:
Amyloglucosidase (AMG), EC 3.2.1.3 (exo-α-1, 4-glucan-glucohydrolase) hydrolyzes maltose to form α-D-glucose. After incubation, the reaction is stopped with NaOH.

Steps 2 and 3 result in an end-point reaction:
Glucose is phosphorylated by ATP in a reaction catalyzed by hexokinase. The glucose-6-phosphate which is formed is oxidized to 6-phosphogluconate by glucose-6-phosphate dehydrogenase. In this same reaction, an equimolar amount of NAD+ is reduced to NADH with a resultant increase in absorbance at 340 nm.

The reaction is illustrated below:



6-phosphoglucc

The change in absorbance is proportional to the enzyme activity within the measuring range; hence a standard linear curve is used.

Reaction conditions

Parameter	Reaction conditions
AMG incubation (step 1)	
Temperature	37.0°C ± 1.0°C
pH	4.30 ± 0.05
Substrate concentration	18 g/L (100 mM) maltose
Enzyme concentration	0.0056 – 0.0167 AGU /ml
Reaction time	360 sec
Color reaction with GHK kit (steps 2 and 3)	
Temperature	37.0°C ± 1.0°C
pH	approx. 7.8
Hexokinase	> 0.66 U/ml
Glucose-6-P-DH	> 0.66 U/ml
Reaction time	420 sec
Wavelength	340 nm
NAD^+	0.74 mM
ATP	0.74 mM

Definition of unit

The amyloglucosidase activity is measured in AGU relative to an enzyme standard of a declared strength.

Method parameters

No significant influence on the analytical results was found for the following additives:
0.6% of Sodium bisulphite
0.2% of Potassium sorbate
0.3% of Sodium benzoate
12% of Sodium chloride
18% of Sorbitol

Equipment

Equipment	
Konelab 30 analyzer	Thermo Electron Corporation
Diluter	E.g., Hamilton Microlab
Analytical balance	E.g., Sartorius, Mettler
Balance	E.g., Sartorius
Magnetic stirrer plates	-

Chemicals

Name	CAS no.	Brand
Triton X-100	9002-93-1	E.g., Sigma T9284
Sodium acetate, 3 H_2O p.a.	6131-90-4	E.g., Merck 6267
Concentrated Acetic acid p.a.	64-19-7	E.g., Merck 63
D-(+)-Maltose monohydrate	6363-53-7	Fluka 63418 lot BCBB7616 V – to be stored at 4°C
NaOH 0.5 M	1310-73-2	Solution, e.g., Titrisol Merck 109956 (Titrisol ampoule for preparation of 1000 ml of 1 M NaOH) or pellets, e.g., Merck 106498
Glucose reagent kit (GHK)	-	Thermo Fisher Scientific, code 981304 or 981779, Glucose HK. To be stored at 4°C. IMPORTANT: Storability of GHK kits for use in this method is 10 months less than the expiration date indicated by the supplier. Do not use GHK kits which are older. Order freshly produced batches (storability according to supplier is approx. 20 months). Upon reception of GHK kits from supplier, a test on each batch must be performed by analyzing 1 standard curve on 3 days of analysis. Test requirements: $r^2 > 0.9985$, slope value ≥ 14.0 Abs /AGU /ml.

Always read the Safety Data Sheet (SDS) for all the chemicals

Reagents

Diluent with Triton X-100, 1 mL/L

EXAMPLE: Preparation of 10 L

Step	Action
1	Weigh out 10.65 g ± 0.01 g of Triton X-100 from Sigma T9284 into a suitable glass beaker. If Sigma T9284 is not used, then make sure to correct the amount in g of Triton X-100 to correspond to 10 ml
2	Add approx. 800 ml deionized water
3	Stir the solution for min. 15 minutes (high speed)
4	Transfer the Triton X-100 solution quantitatively into a 10 L volumetric flask
5	Fill up to the mark with deionized water
6	Stir the final solution for min. 10 minutes
7	Stability: 1 week at room temperature

Acetate Buffer 1.0 M

EXAMPLE: Preparation of 2000 ml

Step	Action
1	Weigh out 88.8 g ± 0.1 g of sodium acetate, 3 H_2O p.a. (e.g. Merck 6267) into a suitable glass beaker
2	Dissolve sodium acetate by addition of approx. 1L deionized water and stir until everything is dissolved. Min. 15 minutes
3	Add 77 ml concentrated acetic acid p.a. (e.g. Merck 63) and stir for min. 5 minutes
4	Transfer the solution to a 2000 ml volumetric flask and fill up to the mark with deionized water
5	Stir for the buffer solution for min. 10 minutes
6	Stability: 1 month at room temperature

Acetate buffer 0.1 M pH 4.30 (AGU-BUF)

EXAMPLE: Preparation of 1000 ml

Step	Action
1	Dispense 100 ml acetate buffer 1.0 M into a suitable glass beaker
2	Add approx. 500 ml deionized water
3	Weigh out 1.07 g ± 0.01 g of Triton X-100 Sigma T9284 and transfer quantitatively to the glass beaker
4	If Sigma T9284 is not used, then make sure to correct the amount in g of Triton X-100 to correspond to 1 ml
5	Stir the solution for min. 15 minutes (high speed)
6	Adjust pH to 4.30 ± 0.05 using NaOH 1 M or HCl 1 M as appropriate
7	Transfer the solution quantitatively from the beaker to a 1L volumetric flask and fill up to the mark with deionized water
8	Stir the buffer solution for min. 5 minutes
9	Check that pH is 4.30 ± 0.05. If not, then adjust pH to 4.30 ± 0.05 using NaOH 1M or HCl 1M as appropriate

Maltose substrate 21.6 g/L (120 mM) (AGU-SUB)

EXAMPLE: Preparation of 100 ml

Step	Action
1	Weigh out 2.16 g of maltose, H_2O p.a. (Sigma 63418) into a 100 ml beaker
2	Add approx. 70 ml acetate buffer 0.1 M pH 4.30 and stir until everything is dissolved. Min. 15 minutes
3	Transfer the solution quantitatively to a 100 ml volumetric flask and fill up to the mark with Acetate buffer 0.1 M pH 4.30. Stir for min. 5 minutes
4	Check that pH is 4.30 ± 0.05. If not, discard the solution
5	Stability: 10 days in refrigerator

NaOH, 0.5 M (AGU-STOP)

EXAMPLE: Preparation of 2000 ml

Step	Action
1	Add the content of one Titrisol ampoule (NaOH solution for preparation of 1000 mL 1 M NaOH) or 40.0 g NaOH pellets to a volumetric flask of 2000 ml
2	Fill up to the mark with deionized water
3	Stir for min. 15 minutes
4	Stability: 2 month at room temperature

Glucose reagent (GHK)

Step	Action
1	Open both reagent A and reagent B bottles (Thermo Fisher Scientific code 981304 or 981779, Glucose HK)
2	Mix the 2 solutions by pouring them in a beaker and then stirring for 2 minutes with a magnetic bar
3	Pour the mix back in reagent A bottle. Check that there are no bubbles in the bottleneck or on the surface of the reagent container and remove or strike out the bar code before you set the bottle into the Konelab analyzer
4	Stability: 30 days in refrigerator

Standard

The standard is available upon request.

Preparation of standards

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>1</td><td>Make sure the standard has reached room temperature</td></tr>
<tr><td>2</td><td>Mix the standard (shake it up and down until it is homogeneously). Make sure there are no brown precipitations in the bottom of the vial</td></tr>
<tr><td>3</td><td>Prepare a stock solution of the standard containing 1.0 AGU/ml with diluent containing Triton X-100.

IMPORTANT: As the standard is a liquid formulation, it is important to reduce the risk of evaporation. The weighing and dissolving of the standard should be carried out as quickly as possible.

NOTE: Open standard vials should be discarded each day. Do not store vials after opening</td></tr>
<tr><td>4</td><td>Stir for min. 15 minutes</td></tr>
<tr><td>5</td><td>Storability of stock solution: 4 hours at room temperature</td></tr>
<tr><td>6</td><td>Working solutions:
Stir the stock solution shortly before diluting to seven levels with diluent containing Triton X-100 directly into sample cups according to the following table.

<table>
<tr><th rowspan="2">Standard no.</th><th colspan="2">Example</th><th rowspan="2">Dilution ratio</th><th rowspan="2">Concentration (AGU/ml)</th></tr>
<tr><th>Stock solution (µl)</th><th>Diluent (µl)</th></tr>
<tr><td>1</td><td>30</td><td>1470</td><td>50</td><td>0.02000</td></tr>
<tr><td>2</td><td>50</td><td>1450</td><td>30</td><td>0.03333</td></tr>
<tr><td>3</td><td>60</td><td>1440</td><td>25</td><td>0.04000</td></tr>
<tr><td>4</td><td>75</td><td>1425</td><td>20</td><td>0.05000</td></tr>
<tr><td>5</td><td>100</td><td>1400</td><td>15</td><td>0.06666</td></tr>
<tr><td>6</td><td>125</td><td>1375</td><td>12</td><td>0.08333</td></tr>
<tr><td>7</td><td>150</td><td>1350</td><td>10</td><td>0.1000</td></tr>
</table>

Storability in sample cups: 40 minutes at room temperature</td></tr>
</table>

Control sample

It is advisable to include a control sample in each run. Prepare the control sample in the same way as described for the samples below always using the same preparation procedure.

Samples

Lowest dilution of all samples (except for BG granulates): 25 ml/g
Lowest dilution of BG granulates: 300 ml/g

Preparation procedure

Step	Action
1	Weigh out and transfer quantitatively to a measuring flask containing diluent containing Triton X-100
2	In the case of BG granulate, fill the measuring flask to roughly the quarter mark and wet the granulate thoroughly. This can be done by rotating the measuring flask slowly until there is no granulate floating on the surface of the diluent containing Triton X-100. IMPORTANT: Do not proceed to the next step before complete wetting is achieved
3	Fill the measuring flask to the mark with diluent containing Triton X-100. Stir for 15 minutes. NOTE: If possible, the activity in the final dilution should be approx. 0.06 AGU/ml
4	Storability of stock solution: Up to 4 hours at room temperature
5	Stir the stock solution prior to dilution
6	Dilute the dissolved samples in diluent containing Triton X-100 directly into sample cups
7	Stability in sample cups: 40 minutes at room temperature

Blank

No reagent blank is used in the method.

Procedure

<table>
<tr><th>Step</th><th>Action</th></tr>
<tr><td>1</td><td>Prepare reagents and dilutions of standard, control sample and samples</td></tr>
<tr><td>2</td><td>Place the reagents in the Konelab:
<table>
<tr><th>Reagent</th><th>Konelab reagent name</th><th>Reagent container volume</th><th>Syringe speed</th></tr>
<tr><td>Acetate buffer (0.1 M pH 4.30)</td><td>AGU-BUF</td><td>20 ml</td><td>Normal</td></tr>
<tr><td>Maltose substrate (21.6 g/L)</td><td>AGU-SUB</td><td>20 ml</td><td>Normal</td></tr>
<tr><td>NaOH 0.5 M</td><td>AGU-STOP</td><td>20 ml</td><td>Normal</td></tr>
<tr><td>Glucose reagent (GHK)</td><td>GHK</td><td>60 ml</td><td>Normal</td></tr>
</table>
NOTE: The solutions in the reagent containers, which have been used in Konelab Arena 30, must be changed every day</td></tr>
<tr><td>3</td><td>Place the standards, control sample and samples in the Konelab in the stated order.
<table>
<tr><th>Position</th><th>Cup with</th></tr>
<tr><td>1-7</td><td>Standard 1-7</td></tr>
<tr><td>8</td><td>Control sample</td></tr>
<tr><td>9-27</td><td>Samples</td></tr>
</table>
19 samples can be analyzed in one analytical run</td></tr>
<tr><td>4</td><td>Start analysis on the Konelab</td></tr>
</table>

Calculation

Step	Action
1	The activity of the enzyme samples is determined relative to the standard curve
2	Based on the results in Abs for the seven AGU standards, a standard curve is drawn with the activities of the standards in AGU/ml as the x-values and the absorbencies of the standards as the y-values using a linear regression model. The concentration of the diluted samples is read from the standard curve using a linear model
3	The activity results are calculated automatically in the AnEx program. Calculation of activity of a sample in AGU/g is performed as stated in the formula: $Activity\ (AGU/g) = \frac{Conc \times V \times F}{W}$ Conc = Reading from the standard curve in AGU/ml V = Volume of the measuring flask used in ml F = Dilution factor for final dilution W = Weight of sample in g
4	EXAMPLE: 0.6098 g sample is dissolved in a 250 ml flask and further diluted 30 times. An Abs of 1.466 is measured on Konelab. From the standard curve an activity of 0.0627 AGU/ml is calculated. $\text{Activity} = \frac{0.0627 \times 250 \times 30}{0.6098} = 771\ \text{AGU/g}$

Approval of analytical run

Standard curve

Parameter	Requirement
Limits for upper and lower y-axis measurements	Absorbance of standard 1 : [0.38 – 0.70] Absorbance of standard 7 : [1.55 – 2.00] If the measured values are not within the approval interval, this might indicate an issue with the maltose substrate solution or the GHK kit.
Residual for upper standard point (standard 7)	The numerical value must not be higher than the residuals for standard 1 – 6. Otherwise, this might indicate an issue with the maltose substrate solution or the GHK kit.
Limits for slope	Slope value: [14.0; 16.9] Abs/AGU/ml. If this is not the case, consider whether dilution of standard is not performed correctly or if the GHK kit has expired.
Quality of fit	$r^2 \geq 0.9985$
Curve appearance	Linear increasing curve

Control sample

It is advisable only to approve runs within ± 2 standard deviations.

Samples

The analytical result (= average of two weighings on three different standard curves) must be CV ≤ 3.6%.

Statement of analysis results

The analytical result is stated with three significant digits.

Configurations

Konelab test definition

```
================================================================================
Test definition              Arena 7.1AR2                          Page:    1

AGU                          Laboratory
                             Analyzer K30/11 - F1120479
Date :   12/06/2009
Time :   10:46
--------------------------------------------------------------------------------

Last change date 12/06/2009 10:46
Tick length (sec) 4.5
Full name           EB-SM-0131
Online Name         AGU                 Test In Use          YES
Test type           Photometric                              LOW     HIGH
                                        Test limit           0.00000 *       Abs
Result unit         Abs                 Initial absorbance   0.000   *       A
Number of Decim.    5                   Dilution limit       *       *       Abs
                                        Secondary dil 1+     0.0     0.0
                                        Critical limit       *       *       Abs
                                        Reflex test limit    *       *       Abs
                                        Reflex test

Acceptance                Automatic     Reference class      LOW     HIGH    In Use
Dilution 1+               0.0

Sample type               Sample type 5 Correction factor    1.00
                                        Correction bias      0.00    Abs
                                        Temperature          37.0    °C

Calibration type          None
Factor                    1.00          Bias                         0.00

Bias correction in use    NO

Manual QC in Use          NO            Routine QC in Use            NO


Blank                     Sample
                                        Normal cuvette

Reagent                   AGU-SUB       Volume (ul)                  100
Disp. with                Extra         Add. Volume (ul)             50
Wash reagent              AGU-BUF       Repl. reagent                AGU-BUF
Reagent wash              Before dispense
Syringe speed             Normal

Incubation                              Time (sec)                   480

Sample                                  Volume (ul)                  20
Disp. with                Extra         Add. Volume (ul)             50
Dilution with             Water         Wash reagent                 None

Incubation                              Time (sec)                   360

Reagent                   AGU-STOP      Volume (ul)                  20
Disp. with                Extra         Add. Volume (ul)             50
Wash reagent              None
Syringe speed             Normal

Incubation                              Time (sec)                   60

Reagent                   GHK           Volume (ul)                  110
Disp. with                Extra         Add. Volume (ul)             50
Wash reagent              None
Syringe speed             Normal

Incubation                              Time (sec)                   420

Measurement
Wavelength
Meas. type
```

AGU-SUB:

Vial vo
Barcode

Syringe

AGU-BUF:

```
==============================================================================
Reagent definition          Konelab Arena 6.5                    Page:    1

                            Novozymes S/N N082 1060
                            Enzym Kemisk Laboratorium


12.10.2007     13:28
------------------------------------------------------------------------------

Reagent                     AGU-BUF        Lot        Expiry date (dd.mm.yy)
                                                     ----------------------------------

Stable on board (days)      1
Alarm limit (ml)            2.0

Information


Vial volume                 20 ml
Barcode id

Syringe speed               Normal
```

AGU-STOP:

```
==============================================================================
Reagent definition          Konelab Arena 6.5                    Page:    1

                            Novozymes S/N N082 1060
                            Enzym Kemisk Laboratorium


12.10.2007     13:28
------------------------------------------------------------------------------

Reagent                     AGU-STOP       Lot        Expiry date (dd.mm.yy)
                                                     ----------------------------------

Stable on board (days)      1
Alarm limit (ml)            1.0

Information


Vial volume                 20 ml
Barcode id

Syringe speed               Normal
```

GHK:

```
==============================================================================
Reagent definition          Arena 7.1AR2                         Page:    1

                            Laboratory
                            Analyzer K30/11 - F1120479


12/06/2009 10:45
------------------------------------------------------------------------------

Reagent                     GHK            Lot        Expiry date (dd/mm/yyyy)
                                                     ----------------------------------

Stable on board (days)      1
Alarm limit (ml)            2.0

Information


Vial volume                 60 ml
Barcode id

Syringe speed               Normal
```

Handling of enzymes and chemicals

Enzymes and enzyme solutions should be handled in a fume hood or in closed containers.
Avoid inappropriate handling of enzymes and enzyme solutions, which may result in aerosol/dust generation.
Avoid inhalation of dust aerosols and contact with skin and eyes.
Handling of chemicals and disposal of waste must be performed according to valid procedures.

Validity

Valid from September 2017.

Novozymes A/S
Krogshoejvej 36
2880 Bagsvaerd
Denmark

Sep 22, 2017 · Luna No. 2010-02878-09

Document:	Method UT.015A	**No.:**	1.9

Title:	Methods for Determination of Elements in Solid and Liquid Enzyme Samples and Samples of Polysaccharides by ICP-MS with Microwave-induced Sample Preparation		**Effective date:**	01.03.2012
			Supersedes:	1.8
			To be revised:	March 2015
Prepared by:	NB	**Approved by:** EVP	**Date:**	01.03.2012

Danish Technological Institute

Chemistry and Microbiology – Taastrup

20.04.2012

Document:	Method UT.015A	No.:	1.9

Area of application

The method is applicable to solid and liquid enzyme samples and samples of polysaccharides for the ICP-MS determination of:

Ag, As, Bi, Cd, Co, Cu, Hg, Mo, Ni, Pb, Sb, Sn and Zn.

The method determines the total content of the specified elements in the stated matrices. The results are used in connection with product control.

The total heavy metal content given by ∑T comprises the elements Ag, As, Bi, Cd, Cu, Hg, Mo, Ni, Pb, Sb and Sn. ∑T is reported as less than the sum of the product limits for the heavy metals set out in Table 1, where these cannot be detected in a sample. In cases where one or more heavy metals are measured at a concentration above the stated product limit(s), the measured value is included in the sum, which will then be a figure greater than 3.0 mg/kg.

The elements in Table 2 are not covered by the accredited method, but are routinely analysed with the accredited elements.

Table 1. **Detection and product limits**

Element	DL accr. no. 90 ppm (mg/kg)	Product limit ppm (mg/kg)	DTI product limit ppm (mg/kg)
As	0.02	1	0.1
Ag	0.01	-	0.5
Bi	0.01	-	0.5
Cd	0.01	0.5	0.05
Co	0.01	-	-
Cu	0.02	1	0.1
Hg	0.01	0.3	0.03
Mo	0.02	-	0.1
Ni	0.04	1	0.1
Pb	0.02	10	1
Sb	0.02	-	0.5
Sn	0.01	-	0.5
Total heavy metal content		40	3.5

Table 2. **Non-accredited elements**

Element	DL ppm (mg/kg)	Product limit ppm (mg/kg)	DTI product limit ppm (mg/kg)
Cr	0.01	-	0.1
Se	0.02	-	0.2
Zn	0.1	-	0.5

Table 3. **Measuring capability**

Measuring capability

Parameter	Upper meas. limit mg/kg	Quanti-sation limit mg/kg	Detection limit mg/kg	%RSD
Ag	5	0.5	0.01	15
As	10	0.1	0.02	15
Bi	5	0.5	0.01	15
Cd	5	0.05	0.01	17
Co	5	0.1	0.02	15
Cu	10	0.1	0.02	30
Hg	5	0.03	0.01	16
Mo	10	0.1	0.02	16
Ni	10	0.1	0.04	22
Pb	100	0.5	0.02	17
Sb	5	0.5	0.003	20
Sn	5	0.5	0.002	20

See annex 2 for uncertainty budgets

Principle

After the dry or liquid product has been weighed out, it is digested with nitric acid in closed Teflon autoclaves by microwave-induced heating or autoclaving. The digested material is then diluted, filtered and analysed by ICP-MS for contents of Ag, As, Bi, Cd, Co, Cu, Hg, Mo, Ni, Pb, Sb, Sn and Zn.

Laboratory equipment

Weighing boats, plastic disposable syringes
Analytical balance ± 0.1 mg
Microwave oven
100-150 ml PFA
Funnels (d = 40 mm with thick stalk)
Filter paper Munktell OOK (d = 110 mm, d = 90 mm)
Volumetric flasks, 50 and 100 ml
Polyethylene vessels, 50 and 100 ml (e.g. Kautex, Nalgene)
Fine pipettes
Autosampler glass, 15 and 50 ml (e.g. Hounisen PP).

Reagents

Demineralised water, Millipore Q-plus
Concentrated nitric acid (HNO_3 14 M) subboiling prepared from Merck par
1.75 M and 2.8 M HNO_3 from concentrated (Suprapur). See Instruction T1801c for preparation.

Standards

Standards of 0, 2, 10 and 50 µg/l of the elements in 2.8 M HNO_3 Prepared as per Instruction T1801c, although standards of 0, 1, 2 and 5 µg/l are used for Hg.

Calibration control

As control solutions, use a "Control I" of 10 µg/l and a "Control II" of 250 µg/l, again prepared as per Instruction T1801c. For Hg use a control Hg_0.1 of 0.1 µg/l prepared as per Instruction T1801c.The operational acceptance criterion for control I, II and Hg_0.1 are within ± 10% of the control values; if this is not met, the person responsible for the analysis should be consulted. Control I and Hg_0.1 are recorded in control charts. See Instruction T 1804.

Control samples

Merck ICP multielement standard solution VI, prepared as the sample, is included as control sample. The operational acceptance criterion for Merck VI is ± 10 % of the control values; if this is not met, the person responsible for the analysis should be consulted.

Merck VI is further recorded in control charts. See Instruction T 1804.

Sample handling

Solid samples are stored at room temperature.

Liquid samples are stored in a refrigerator at < 5°C until the time of analysis.

Samples received frozen are stored in a freezer until the time of analysis, then defrosted in a refrigerator.

Preparing equipment

Cleaning

The PFA autoclaves are cleaned with 20 ml conc. HNO_3 per vessel by microwaving for 20 minutes at 100% and then rinsing with demineralised water.

Preparing samples

Solid samples

0.5 ± 0.1 g of sample is weighed out to 4 decimal places in a weighing boat and transferred to a PFA autoclave, after which the weigh boat is reweighed. The sample is suspended in 20 ml 7 M HNO_3.

The autoclave is closed, placed in the carrousel and microwaved for 35 minutes at 630 W with regulation to max. 230 °C and 20 bar in accordance with the instructions for use of the microwave oven.

The autoclave is opened after it has cooled.

The contents are filtered into a 50 ml acid-rinsed volumetric flask with demineralised water and diluted to volume with demineralised water. The filtrate is stored in a polyethylene vessel.

Liquid samples

1 ± 0.5 g of sample is weighed out to 4 decimal places in a weighing boat and transferred to a PFA autoclave, after which the weighing boat is reweighed. When weighing out highly viscous enzyme samples, a disposable syringe serves well as a weighing bottle.
The sample is suspended in 20 ml 7 M HNO_3.
The autoclave is closed, placed in the carrousel and microwaved for 35 minutes at 630 W with regulation to max. 230°C and 20 bar in accordance with the instructions for use of the microwave oven.

The autoclave is opened after it has cooled.

The contents are filtered into a 50 ml volumetric flask with demineralised water and diluted to volume with demineralised water. The filtrate is stored in a polyethylene vessel.

The sample preparation for heavy metals (Ag, As, Bi, Cd, Co, Cu, Hg, Mo, Ni, Pb, Sb, Sn and Zn) are carried out as duplicate determinations.

At least 2 blind samples are prepared accordingly for each sample series.

Analysis

The analysis is performed by FI-ICP-MS (flow injection inductively coupled plasma mass spectrometry) using external standards in 2.8 M HNO_3, and adding the internal standards in 0.14 M HNO_3 to all measuring solutions online via a flow injection system. Typical instrument parameters are given in Tables 4 and 5.

A calibration blank of 2.8 M HNO_3 is prepared.

The system is rinsed with 1.75 M HNO_3 (carrier).

The autosampler probe is rinsed with 1.75 M HNO_3 (from Subboiling).

An example of a routine analysis set-up is shown in Annex 1. This will typically include the calibration blank, a standard series, detection limit standards at 2 levels, e.g. 1 and 5 ng/ml, Controls I and II, an Hg control, and a double-determination on a control sample (Merch VI) that is entered on the control chart.

After every 10 – 12 samples, or on completion of a run, a calibration blank and controls I and II are analysed.

The controls should not deviate more than stated under 'calibration control'.

See Table 1 for detection limits.

An automatic dilution factor should never be used.

Generally, the sample is diluted at concentrations of the analysis subject greater than 500 ng/ml. Alternatively, the sample is re-analysed by ICP-AES for these parameters.

A toxbatch sample is reanalysed, if %RSD is larger than 20 % for a double determination of an element, when the concentration is above the quantification limit of the method. All valid subresults are reported. This applies only for toxbatch/GLP-samples.

Table 4. **Spectrometer parameters**

Element	Mass	Dwell time	Notes
Ag	107	17	
As	75	17	Corr (Se)
Bi	209	17	
C	13	17	-
Ca	43	17	-
Cd	111	17	
Co	59	17	Corr (Ca)
Cr	52	17	*, Corr (C)
Cu	65	17	
Ge	72	17	IS
Hg	202	17	Corr (W)
Mo	98	17	
Ni	60	17	
Pb	208	17	
Re	187	17	IS
Rh	103	17	IS
Sb	121	17	
Se	77	17	-
Se	82	17	*
Sn	120	17	
Tl	205	17	*
W	182	17	-
Zn	66	17	*

IS denotes internal standards.
Corr(X) denotes that the measurement is corrected via element X.
* denotes that elements are not covered by the accreditation.
- denotes that the given element is used for a correction.

Table 5. **Plasma parameters**

Plasma flow (l/min)	15
Neb. flow (l/min)	0.6-1.2
Aux. flow (l/min)	0.8
RF power (W)	1100-1200
Perimax 16 (rpm)	-11
Replicates	3

Calculations

Data processing and calculation of results

Data from ICP-MS is transferred to an external computer for the results to be calculated using the LISA calculation program.

Calculation principle:

The concentration in the sample, C_{sample}, is calculated as follows:

$$C_{sample} = (V \cdot (F \cdot C_{measured} - C_{blank}))/m_{sample}$$

where

$C_{measured}$ is the concentration in the measuring solution
C_{blank} is the concentration in the blind sample. If C_{blank} is < the detection limit the value 0 is used
F is the dilution factor (normally 1)
V is 50 ml
m_{sample} is the quantity of sample weighed-out.

If the result is given based on dry matter, C_{sample} must be corrected for percentage dry matter (%DM).

$$C_{sample} = (100 \times C_{sample})/\%DM$$

Reporting

Reports on liquid and solid enzymes are made directly on the analysis request forms and returned by fax. The following information is added to the request form:

- date of receipt/initials
- The Laboratory for Chemistry and Microbiology's sample ID
- date/initials
- results for the parameters selected by the person placing the order.

After mailing, the job is considered reported.

In the case of tox batch samples (GLP jobs), an authorised written report of the analysis is sent with full documentation in the form of an annex.

The annexes comprise:

1) Registration slip
2) Weighing chart with weight ID, forms for preparation of intermediate dilutions, standards incl. standards for Hg and for control I and II, Sample Batch Report, Dataset Report, calculations and raw data.
   On the first page of the printed raw data, the instrument ID is specified.

No. T-1.0012 for PE Sciex Elan 5000 and T-1.0018 for PE Sciex Elan 6100, respectively. The raw data sheets are clipped together.

3) Instrument parameters

Safety

Nitric acid HNO_3 conc.

Hazard symbols: O (oxidising) + C (corrosive).

R/S phrases: R8-35 S 23-26-36.
Causes burns.
Contact with combustible material may cause fire.
Irritating to respiratory system.
In case of contact with eyes, rinse immediately with plenty of water and seek medical advice.
Wear suitable protective clothing and protective gloves.
In case of accident or if you feel unwell, seek medical advice immediately (show the lable where possible).

Spills: Absorb with ABSOL. Strong ventilation.

Precautions: Work with nitric acid should be carried out in a fume hood, wearing gloves and safety goggles.
See directions for use.

When weighing enzyme samples, especially powder, wear a mask and safety gloves. See reference 5 on handling enzyme samples.

Literature

1. Determination of Trace Elements in Waters and Wastes by Inductively Coupled Plasma Mass Spectrometry. U.S. Environmental Protection Agency. Method 200.8, Revision 4.4, April 1991.

2. Users Manual
ELAN 5000 Inductively Coupled Plasma Mass Spectrometer
Perkin Elmer, May 1992, rev. B

3. Users Manual
ELAN 6100 DRC Inductively Coupled Plasma Mass Spectrometer Perkin Elmer, May 2000

4. Report 17661, 21.12.1995:
"Comparative analysis of 9 granulated enzyme samples and of 11 liquid enzyme samples for element contents by FI-ICP-MS and ICP-AES".

5. "Enzymes and you – a guide for laboratory workers".
Novo Nordisk A/S.

Annex 1 Sample/Batch Report

Annex 2 Uncertainty budgets

Document:	Method UT.015A	No.:	1.9

Correction log

Correction page no	Previous version	New version
2	Mo in table 1	Mo removed from table 1
3	Zn missing in table 2	Zn added to table 2
6	A single preparation is made for heavy metals (Ag, As, Bi, Cd, Co, Cu, Hg, Mo, Ni, Pb, Sb, Sn and Zn), although double determinations are always carried out for tox batch samples (GLP analyses).	"Analysen for tungmetaller (Ag, As, Bi, Cd, Co, Cu, Hg, Mo, Ni, Pb, Sb, Sn og Zn) udføres som dobbeltbestemmelse".
6	In cases where one or more heavy metals are outside the product limit (see Table 1), the analysis is repeated with a double preparation.	The text is deleted.
7	Litt. Ref. includes 4 references.	Added as item 3 : Users Manual ELAN 6100 DRC Inductively Coupled Plasma Mass Spectrometer Perkin Elmer, May 2000. Previous item 3. was changed to item 4. and previous item 4. was changed to item 5. Text unchanged.

Enumeration of Total Viable Count

Scope

All Novozymes QC laboratories involved in analysis of samples from Novozymes production and GLP studies.

Principle

Total Viable Count (TVC) is defined as the number of organisms which form colonies on a non-selective agar medium (Tryptic Soy Agar, TSA) after aerobic incubation for 3 days at 30-35°C. TSA is a rich non-selective agar medium.

The method outlined below conforms to the principles of The European Phamacopoeia (EP), Chapter 2.6.12 B. Harmonised method with the following exceptions:

- The test only covers the enumeration of microorganisms capable of growing on TSA (Total aerobic Microbial Count). The test does not include enumeration of Total combined Yeast / Mould Count.
- The dilution water has an addition of 4% Tween 80.
- EP describes the use of duplicates. This method uses single tests.
- The agar plates are incubated for 3 days, not for 3-5 days.
- Growth promotion test of TSA is performed according to in-house procedures and not according to the description in EP.

Samples are analysed by the spiral plater (100 µl) or spread plate technique (100 µl or 1 ml) from suitable dilutions:

Sample type	Requested test (LIMS code)	Technique	Volume spread	Lowest Dilution	No. of plates	Plate size	Detection limit
Enzyme samples and fluid hyaluronic acid	TVC	Spiral plating or spread plating	100 µl	10^{1}	1 plate	9 cm	100 CFU / g or ml
	TVC(100)	Spread plating	1 ml	10^{1}	4 plates	14 cm	10 CFU / g or ml
CIP samples	CIP_TVC	Spiral plating or spread plating	100 µl	Undiluted	1 plate	9 cm	10 CFU / ml
FeF samples	FEF_TVC	Spread plating	1 ml	10^{1}	4 plates	14 cm	10 CFU / g or ml
Intermediates	TVC FAST_xxxx	See EB SM 3200.02 D					
Swabs	SWAB_ TVC	See EB SM 5001.02					
Water samples	WATER_ TVC (xx xx)	See EB SM 3095.02 D					

Continued on next page

Enumeration of Total Viable Count, *Continued*

Definition of units

The result is stated as:

- Total Viable Count (TVC) / g or ml

Samples

All sample types.

Detection limit

The detection limit of this method is dependent on the sample volume and the dilution in use (See “Principle”).

Equipment

Balance (± 0.1 g)
Magnetic stirrer
Petri dishes (9 cm or 14 cm)
Suitable sterile pipettes for transfer of 100 µl or 1 ml (4x0.25 ml)
Spiral plater (for the spiral plate technique)
Sterile Drigalski spreaders (for the spread plate technique)
Incubator (30-35°C)

Media and reagents

Tween buffer 4%, 90 ml (if necessary with a magnet) prepared acc. to EB-ME-0052
EP buffer, 90 ml buffered sodium chloride-peptone solution pH 7.0, prepared acc. to EB-ME-0067
TSA plates (9 or 14 cm) prepared acc. to EB-ME-0041

Safety

It is the responsibility of the laboratory leader, that all personnel are aware of the correct handling of enzymes and reagents.

Continued on next page

Enumeration of Total Viable Count, *Continued*

Sample preparation

Enzyme samples and other solid samples are prepared as follows:

Sample type	Action
Enzyme samples **FeF samples** **Other solid samples**	Transfer 10 g of solid sample or 10 ml of liquid sample to 90 ml Tween buffer 4%. *NOTE:* Immediately homogenize the sample by stirring or by shaking. Solid samples are homogenized on a magnetic stirrer for app. 20 minutes. *TIP:* Further 10-fold dilutions can be prepared with Tween buffer 4%.
Non-enzyme fluid samples (e.g. CIP samples)	Fluid samples are analyzed undiluted. *TIP:* If needed, 10-fold dilutions may be prepared with Tween buffer 4%.
Fluid hyaluronic acid (HA)	Transfer 10 ml of liquid sample to 90 ml EP buffer. *IMPORTANT:* Homogenize on a magnetic stirrer for min. 20 minutes and max. 1 hour.
Air monitoring samples	Agar plates are incubated when received in the laboratory.

IMPORTANT: All enzyme products must be analyzed from a 10^{-1} dilution due to possible inhibition of microorganisms in undiluted enzyme.

Non-enzyme liquid samples (e.g. CIP-samples) are analyzed undiluted, if relevant.

Plating

Plating must be done within 15 minutes from end of homogenisation. If this is not possible, the sample can be stored at 2-8°C for up to 4 hours.

Test	Action
TVC	Transfer 100 µl from the 10^{-1} dilution onto the surface of a TSA plate (9 cm). Repeat this for any of the necessary dilutions. *or* Perform a spiral plating of 100 µl from the 10^{-1} dilution in accordance with the directions for the specific spiral plater.
TVC(100) or TVC_FeF	Transfer 1 ml from the 10^{-1} dilution onto the surface of 4 TSA plates (14 cm) with app. 0.25 ml onto each plate. Repeat this for any of the necessary dilutions.

TVC_CIP	Transfer 100 µl from the undiluted sample onto the surface of a TSA plate (9 cm). Repeat this for any of the necessary dilutions *or* Perform a spiral plating of 100 µl from the undiluted sample in accordance with the directions for the specific spiral plater.

Leave the plates on the table with lid on until the sample has been soaked into the agar.

Incubation

Incubate the plates at 30-35°C for 3 days.

Continued on next page

Enumeration of Total Viable Count, *Continued*

Reading

Spread plate technique:

Count the number of colonies on the plates.

Size of agar plate	Count colonies on plates with
9 cm	1–300 colonies per plate
14 cm	1–750 colonies per plate

Spiral plate technique:

The number of typical colonies on each plate is counted and the result is calculated in accordance with the directions for the specific spiral plater.

Calculation

General principles:

The calculation is based on the number of colonies (C_x) on the plate, and the sample volume analysed (V_x).

The result is stated with two significant figures (e.g. 2.2×10^1).

When Using results from	Then the result is	Where
One dilution	$\frac{C_x}{V_x}$	C_x = no. of colonies V_x = volume analysed
2 or more dilutions	$\frac{C_1 + C_2}{V_1 + V_2}$	C_1 = no. of colonies in lowest dilution C_2 = no. of colonies in next dilution V_1 = volume analyzed in lowest dilution V_2 = volume analyzed in next dilution

IMPORTANT: When using more than one dilution, the numbers from each dilution are compared (the likelihood of product inhibitions, contamination of the sample, analytical errors etc. is considered). In general, the highest dilution is used. If the result is stated on the basis of other dilutions, the reason must be given in the raw data.

When the sample volume is 0.1 ml then V_x and C_x are:

Dilution	Undiluted	10^{-1}	10^{-2}
V_x	0.1 ml	0.01 ml	0.001 ml
C_x	No. of colonies on the plate	No. of colonies on the plate	No. of colonies on the plate

Continued on next page

Enumeration of Total Viable Count, *Continued*

Calculation *(continued)*

EXAMPLE: Examples of calculating spread plate of 0.1 ml sample:

C_x	V_x (g or ml)	Dilution	Result
0	0.01	10^{-1}	< 100 / g or ml
123	0.1	10^{-0}	$\frac{123}{0.1} = 1.2 \times 10^3$ / g or ml
334	0.01	10^{-1}	> 3.0×10^4 / g or ml
253 24	0.01 0.001	10^{-1} 10^{-2}	$\frac{253+24}{0.01+0.001} = 2.5 \times 10^4$ / g or ml

When the sample volume is 1 ml (four 14 cm agar plates with 0.25 ml on each plate) then V_x and C_x are:

Dilution	Undiluted	10^{-1}	10^{-2}
V_x	1 ml	0.1 ml	0.01 ml
C_x	sum of colonies on the 4 plates	sum of colonies on the 4 plates	sum of colonies on the 4 plates

EXAMPLE: Examples of calculating spread plate of 1 ml sample:

C_x	V_x (g or ml)	Dilution	Result
0	0.1	10^{-1}	< 10 / g or ml
123	1	10^{-0}	$\frac{123}{1} = 1.2 \times 10^2$ / g or ml
426	0.1	10^{-1}	$\frac{426}{0.1} = 4.3 \times 10^3$ / g or ml
3134	0.1	10^{-1}	> 3.0×10^4 / g or ml
853 84	0.1 0.01	10^{-1} 10^{-2}	$\frac{853+84}{0.1+0.01} = 8.5 \times 10^3$ / g or ml

Continued on next page

Enumeration of Total Viable Count, *Continued*

Accuracy and precision	CV% (surface plating) = 25% CV% (spiral plating) = 29% *REFERENCE:* http://luna.novo.dk/cgi-bin/lunaSearch.pl?action=details&luna=2003-34435&db=LUNA
Filing	All documentation should be archived in accordance with the local filing SOP.
Contingencies	All deviations from this SOP should be discussed with the Method Responsible Scientist and should be documented.
References	European Pharmacopoeia, Chapter 2.6.12. Microbiological examination of non-sterile products (Total viable aerobic count).
Revision	Urgent revision of version 6.0, which was never validly published, as BioBusiness by mistake asked for removal of TVC analysis of hyaluronic acid. Hyperlink to swab test updated. Added EP buffer in the section “Media and reagents”. Added the samples types FeF samples, CIP samples, fluid hyaluronic samples, and air monitoring samples in the section “Sample preparation”. Other minor editorial changes.

Enumeration of coliform bacteria using Violet Red Bile agar

IMPORTANT This method is used for the analysis of **Sweetzyme** (batch code IA) and **liquid products** (with the exception of Biofeed Plus, batch code CN).

Scope All Novozymes QC laboratories involved in analysis of samples from Novozymes production and GLP studies.

Principle **Coliform bacteria** are broadly defined as Gram-negative, oxidase-negative, non-sporogenous rods, which grow in aerobic or facultative anaerobic conditions. More specifically, coliforms are capable of fermenting lactose (due to production of galactosidase) in the presence of bile at 37°C. Coliforms are not a taxonomically defined group of bacteria and consequently there is not a common agreement of which microorganisms truly belong to the coliform bacteria. However, as defined in (Ref. 3), Novozymes define coliform bacteria as organisms belonging to the genera *Escherchia*, *Citrobacter*, *Enterobacter, Klebsiella, Serratia* and *Hafnia*. The presence of coliform bacteria, especially *E. coli*, can be used as an indicator of the bacteriological hygiene of an enzyme product.

The Violet Red Bile agar (VRB) is a selective and indicative agar:

Principle	Description
Selective principle	Crystal violet and bile salts inhibit growth primarily of the Gram-positive accompanying flora. This favors growth of the fast growing Gram-negative enterobacteria.
Indicative principle	Degradation of lactose to acid is indicated by the pH indicator neutral red, which changes its color to red and in some cases also by precipitation of bile acids. Coliform bacteria degrade lactose.

Routine testing is performed in the following way:

Sample type	Requested test (LIMS code)	Technique	Volume spread	Lowest dilution	No. of plates	Plate size	Detection limit
Enzyme samples	COLIFORM	Pour plate with cover layer	2½ ml	10 [1]	1 plate	14 cm	4 CFU / g or ml
CIP and water samples	CIP_ COLIFORM	Pour plate with cover layer	1 ml	Undiluted	1 plate	9 cm	1 CFU / g or ml

Depending on sample type, level of contamination and the detection limit needed for the specific sample, alternative procedures may be used.

Continued on next page

Enumeration of coliform bacteria using Violet Red Bile agar, *Continued*

Principle *(continued)*

The method outlined below conforms to ISO 4832 with the following deviations:

- ISO 4832 and ISO 6887-1 describe the use of a Peptone-salt-solution or Buffered-peptone-water as diluent. This Novozymes method uses Tween buffer 4%.
- ISO 4832 describes the use of duplicates. This Novozymes method uses single tests.

Definition of units

The result is stated as:

- Coliform bacteria / g or ml

Samples

This method is used for the analysis of **Sweetzyme** (batch code IA) and **liquid products** (with the exception of Biofeed Plus, batch code CN). Biofeed Plus is analyzed according to EB-SM-3005.

NOTE: In addition, the method can be used for analysis of certain solid samples (e.g. cryst. conc. T).

Detection limit

The detection limit of this method is dependent on the sample volume and the dilution in use (See the section “Principle”).

Equipment

Balance (± 0.1 g)
Magnetic stirrer
Petri dishes (9 cm or 14 cm)
Suitable sterile pipette for transfer of 1 ml or 10 ml (2.5 ml)
Incubator (34-38°C)

Media and reagents

- Tween buffer 4%, 90 ml (If necessary, with a magnet) prepared acc. to EB-ME-0052.
- Violet Red Bile agar (VRB) prepared acc. to EB-ME-0051.

NOTE: If the agar is freshly prepared in the laboratory, suspend the media with 200 ml exchanged water and leave for 15 min. Ensure that media is thoroughly dissolved before melting procedure by regular shaking of the media. In addition, stir the agar immediately before cooling in water bath and again before pouring in Petri dishes.

Continued on next page

Enumeration of coliform bacteria using Violet Red Bile agar, *Continued*

Safety

It is the responsibility of the laboratory leader, that all personnel are aware of the correct handling of enzymes and reagents.

Suspension of VRB agar in water should be executed in a clean bench to avoid inhalation of VRB powder.

Sample preparation

Enzyme samples are prepared as follows:

Step	Action
1	Transfer 10 g of solid sample or 10 ml of liquid sample to 90 ml Tween buffer 4%.
2	Immediately homogenize the sample by stirring or by shaking. Solid samples are homogenized on a magnetic stirrer for app. 20 min.

IMPORTANT: All enzyme products must be analyzed from a 10^{-1} dilution due to possible inhibition of microorganisms in undiluted enzyme.

Non-enzyme liquid samples (e.g. CIP-samples) are analyzed undiluted.

TIP: Further 10-fold dilutions of any sample type can be prepared with Tween buffer 4%.

Plating

Plating is performed using the pour plate technique:

Sample type	Description
Enzymes	1. Transfer 2½ ml from the 10^{-1} dilution to an empty Petri dish (14 cm). 2. Pour app. 40-45 ml VRB agar (47 ± 2°C) in the Petri dish (= bottom layer) and mix carefully. Leave this to solidify. 3. Pour app. 10 ml VRB agar (47 ± 2°C) onto the bottom layer (= covering layer). Leave this to solidify.
CIP and water samples	1. Transfer 1 ml from the undiluted sample to an empty Petri dish (9 cm). 2. Pour app. 20-25 ml VRB agar (47 ± 2°C) in the Petri dish (= bottom layer) and mix carefully. Leave this to solidify. 3. Pour app. 5 ml VRB agar (47 ± 2°C) onto the bottom layer (=covering layer). Leave this to solidify

IMPORTANT: Agar used for BB samples must be cooled to 45 ± 2°C.

Continued on next page

Enumeration of coliform bacteria using Violet Red Bile agar, *Continued*

Incubation

Incubate the plates at 34-38°C (Target = 36°C) for 1 day at aerobic conditions.

Reading

Count the number of typical colonies:

Size of agar plate	Count colonies on plates with	Typical colonies
9 cm	1-150 per plate	Purplish red with a diameter of ≥ 0.5 mm and sometimes surrounded by a reddish zone of precipitated bile.
14 cm	1-375 colonies per plate	

Calculation

General principles:

The calculation is based on the number of colonies (C_x) on the plate, and the sample volume analyzed (V_x).

The result is stated with two significant figures (e.g. 2.2×10^1).

When Using results from	Then the result is	Where
One dilution	$\frac{C_x}{V_x}$	C_x = no. of colonies V_x = volume analyzed
2 or more dilutions	$\frac{C_1 + C_2}{V_1 + V_2}$	C_1 = no. of colonies in lowest dilution C_2 = no. of colonies in next dilution V_1 = volume analyzed in lowest dilution V_2 = volume analyzed in next dilution

IMPORTANT: When using more than one dilution, the numbers from each dilution are compared (the likelihood of product inhibitions, contamination of the sample, analytical errors etc. is considered). In general, the highest dilution is used. If the result is stated on the basis of other dilutions, the reason must be given in the raw data.

When the sample volume is 2½ ml then V_x and C_x are:

Dilution	10^{-1}	10^{-2}
V_x	0.25 ml	0.025 ml
C_x	No. of colonies on the plate	No. of colonies on the plate

Continued on next page

Enumeration of coliform bacteria using Violet Red Bile agar, *Continued*

Calculation *(continued)*

EXAMPLE: Examples of calculating pour plate of 2½ ml sample on a 14 cm agar plate:

C_x	V_x (g or ml)	Dilution	Result
0	0.25	10^{-1}	$\frac{0}{0.25}$ = < 4 / g or ml (LIMS = < 10)
1	0.25	10^{-1}	$\frac{1}{0.25}$ = 4 / g or ml
3	0.25	10^{-1}	$\frac{3}{0.25}$ = 12/ g or ml
412	0.25	10^{-1}	$\frac{375}{0.25}$ = > 1.5 x 10^3 / g or ml
53 8	0.25 0.025	10^{-1} 10^{-2}	$\frac{53+8}{0.25+0.025}$ = 2.2 x 10^2 / g or ml

EXAMPLE: Examples of calculating pour plate of 1 ml sample on a 9 cm agar plate:

C_x	V_x (g or ml)	Dilution	Result
0	1	10^0	$\frac{0}{1}$ = < 1 / g or ml (LIMS = < 10)
18	1	10^0	$\frac{18}{1}$ = 18 / g or ml

IMPORTANT: When the result entered in LIMS is a 'less than' value lower than < 10 / g or ml, LIMS will automatically change this value to "< 10".

Accuracy and precision

CV% = 29%

REFERENCE: LUNA no. 2003-34435

Filing

All documentation should be filed in accordance with the local filing SOP.

Contingencies

All deviations from this SOP should be discussed with the Method Responsible Scientist and should be documented.

Continued on next page

Enumeration of coliform bacteria using Violet Red Bile agar, *Continued*

References

1. ISO 4832 2nd Ed. (1991) Microbiology – General Guidelines for the enumeration of coliforms – colony count technique.
2. ISO 6887-1 1st Ed. (1999) Microbiology of food and animal feeding stuffs - Preparation of test samples, initial suspensions and decimal dilutions for microbiological examination – Part 1: General rules for the preparation of the initial suspension and decimal dilutions.
3. LUNA no. 2009-26425-01: Definition of enterobacteria and coliform bacteria at Novozymes.

Revision

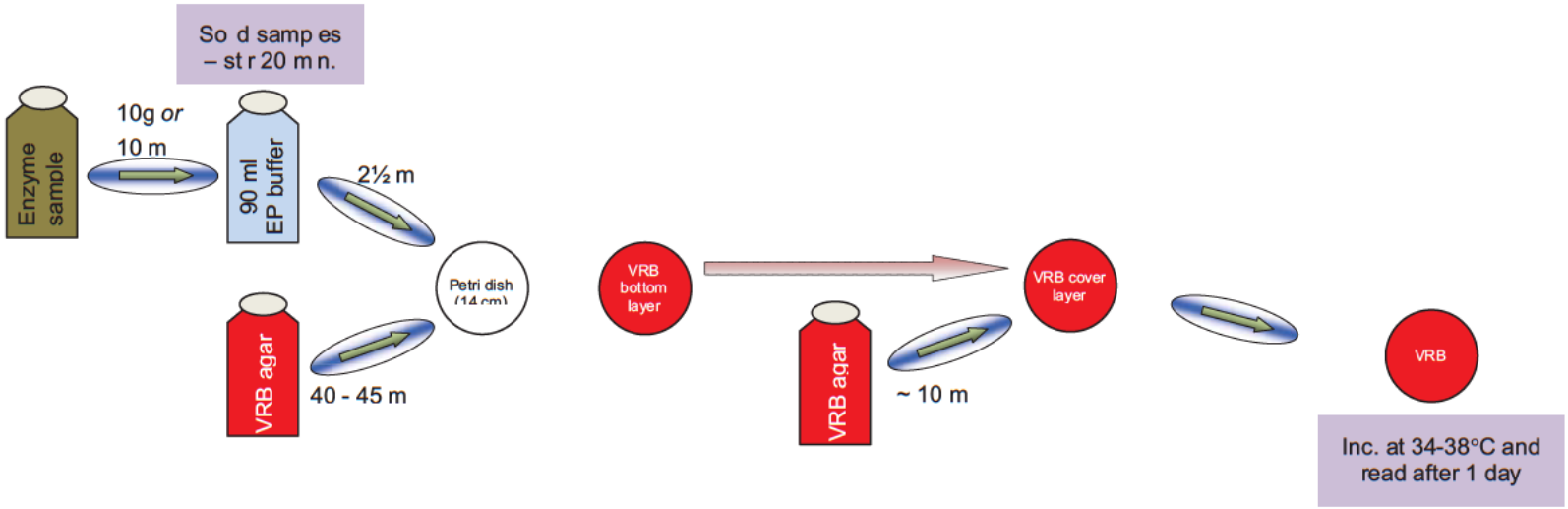
CIP and water samples are analysed using the same procedure. Therefore both sample types are stated in the sections “Plating” and “Principle”. In the section “Calculation” it is emphasized that if results are ‘less than’ values lower than < 10 / g or ml, LIMS will automatically change these values to “< 10”. For instance, when “<4” is entered because 0 colonies were counted when analysing an enzyme sample, LIMS will automatically change this to “<10”. The reason being that “<10” is the only less than value allowed in EB-SM-3005 which is also used for the COLIFORM LIMS method.

Continued on next page

Enumeration of coliform bacteria using Violet Red Bile agar, *Continued*

Appendix – Flow chart

Flow chart of method. Ctrl + Click to read section.



Detection of E.coli in 25 g

Scope

All Novozymes QC laboratories involved in analysis of samples from Novozymes production and GLP studies.

Overview

Principle

Escherichia coli (*E.coli*) is a Gram-negative, indole positive, facultative anaerobic rod. It is considered a faecal indicator.

Detection of *E.coli* in 25 g is carried out as a qualitative analysis using non-selective enrichment in Buffered Peptone Water (BPW) followed by isolation of β-D-glucoronidase positive *E.coli* on a selective indicative agar medium (TBX agar). β-Glucuronidase-negative *E. coli* strains (3-4 %) form colourless colonies on TBX agar, e.g. *E. coli 0157*. The detection of *E.coli O157* is performed as ImmunoMagnetic Separation (IMS) using Dynabeads®antiO157 and plating onto two selective indicative agar media (CT-SMAC agar and ChromAgar O157). Suspect *E.coli O157* colonies are verified using *E.coli O157* Latex test.

Suspect colonies from TBX agar and/or *E.coli* O157 Latex positive isolates from CT-SMAC agar and/or ChromAgar O157 are reported as *E. coli* Detected in 25 g

Suspect colonies are further verified for Enterovirulent *E. coli* (EEC*)* in department 402 Bagsværd (PSL-SM-3097)

Requested test (LIMS code): E.COLI(25g)

IMPORTANT: E. coli (25) analysisis always requested together with Enterovirulent *E. coli* (EEC*)*verification, Lims code: EV E.coli. If E. coli(25g) is Not Detected (ND) both metods are reported as ND in LIMS.

Continued on next page

Detection of E.coli in 25 g, *Continued*

Principle *(continued)*

The media used has the following characteristic:

Media…	Characteristic…
BPW broth	Non-selective broth.
TBX agar	**Selective properties:** Growth of accompanying Gram-positive flora is largely inhibited by the use of bile salts. **Indicative properties:** The presence of the enzyme β-D-glucuronidase differentiates most *E.coli* spp. from other coliforms. *E.coli* absorbs the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-β-D-glucuronide). The enzyme β-glucuronidase splits the bond between the chromophore 5-bromo-4-chloro-3-indolyle- and the β-D-glucuronide. *E.coli* colonies are coloured blue-green. *NOTE:* For the recovery of sublethally injured *E. coli*, plates are incubated at 34 - 38°C and not 44°C as recommended by Merck (inhibits growth of accompanying Gram-positive flora).
CT- SMAC agar (MacConkey Sorbitol agar)	Polypeptone favors the growth of Escherichia coli O157:H7. Sorbitol negative bacterial (in particular O157:H7) colonies are colorless. Sorbitol positive bacteria give rise to red colonies, due to the change of the color of the pHindicator (neutral red). Contaminating bacteria are inhibited by the association of bile salts, crystal violet, cefixime and potassium tellurite**.**
ChromAgar O157 and CT-Chrom Agar O157	A typical *E.coli O157* will grow as a pink-mauve colony, whereas most other micro organisms are either inhibited or grow as blue or colourless colonies.

Definition of units

The result is stated as:

- DET (*E.coli* Detected in 25 g) *or*
- ND (*E.coli* Not Detected in 25 g)

Samples

All Novozymes sample types.

Continued on next page

Detection of E.coli in 25 g, *Continued*

Standards

A positive reference strain can be used, e.g. *E.coli*, ATCC 11229.
If a reference strain of *E.coli O157* is included, it must be *E.coli O157* without the genes coding for Vero Toxins, e.g. ATCC 43888

Detection limit

Theoretical detection limit: 1 *E.coli* in 25 g

Equipment

Balance (± 0.1 g)
Magnetic stirrer
Incubator (34-38°C)
Sterile inoculation loops (1 µl)
Sterile swabs
Vortex mixer
Pipettes and sterile tips
For ImmunoMagnetic Separation (either mIMS or aIMS):

- For manual ImmunoMagnetic Separation (mIMS):
  - MPC-S Rack and magnet (Invitrogen Cat. No. 120.20) + Eppendorf tubes 1.5 ml (Eppendorf Cat. No. 0030 10.086) + MX-3 Mixer (Dynal Cat. No. 159.09), mixer is optional.

- For automatic ImmunoMagnetic Separation (aIMS):
  - BeadRetriever (Invitrogen Cat. No. 159-50) + Tubes & tips (Invitrogen Cat. No. 150-51)

Media and reagents

Buffered Peptone Water (BPW) (450 ml) acc. to EB-ME-0009.
Chromocult ®TBX agar plates (9 cm) acc. to EB-ME-0012
Cefixime-Tellurite-Sorbitol MacConkey agar (CT-SMAC agar plates. 9 cm)
ChromAgar O157 agar plates (9 cm) or CT-ChromAgar O157 (app. 5 or 9 cm)
Tryptone Soya agar plates (TSA)
Dynabeads®anti O157, Dynal Cat. 710.04
Washing buffer (PBS-Tween 20 buffer), Sigma No. P-3563
E.coli O157 Latex test kit (for verification), Oxoid No. DR620
IMPORTANT: Preparation in the local laboratory shall be done according to the current valid WW Media direction.

Continued on next page

Detection of E.coli in 25 g, *Continued*

Safety

- It is the responsibility of the laboratory leader that all personnel are aware of the correct handling of enzymes and reagents.
- E. coli O157 Latex test (Oxoid DR0620) is labelled R22- Harmful if swallowed due to 0.1% Sodium azide.

Transfer of sample to BPW

• **If the sample is …**	**Then transfer 25 g sample to …**
Novamyl SM30 conc. BG (AB…..) Neutrase 1.5 Unstd M (PW…..) Ultraflo Unstd MG (CN …..) Clear Lens Pro 2.5 MG (P…..) Viscoflow MG (KR …..) Flavourzyme 500 MG (HP …..) Ceremix Plus MG (WD …..)	900 ml BPW (Use two 450 ml BPW bottles and transfer 12.5 g to each bottle)
… any other sample	450 ml BPW

Enrichment

The non-selective enrichment is performed in the following way:

- Incubate BPW at 34-38°C for 16-20 hours (minimum 16 hours).

Detection of β-D-glucoronidase positive E.coli

Detection of β-D-glucoronidase positive *E.coli* is performed in the following way:

- Streak the enriched sample onto the surface of a TBX agar plate using a sterile 10 µl inoculation loop. If 2 BPW bottles streak on 1 agar plate from each bottle.
- Incubate the plate at 34-38°C for 18 - 24 hours.
- Examine the plate for growth of typical E.coli colonies:

Organism	Growth on Chromocult®TBX agar
E.coli	Blue-green or dark-blue to violet colonies colonies (Salmon-GAL and X-glucuronide reaction)
Coliforms (not *E. coli*)	Salmon to red colonies (Salmon-GAL reaction but no X-glucuronide reaction)
Other Gram-negatives	Colourless colonies, except for some organisms which possess ß-D-glucuronidase activity. These colonies appear light-blue to turquoise.

Continued on next page

Detection of E.coli in 25 g, *Continued*

Detection of *E.coli O157*

ImmunoMagnetic Separation (IMS) is performed either as manual IMS (= mIMS) or as automated IMS (= aIMS):

Manual IMS (= mIMS):

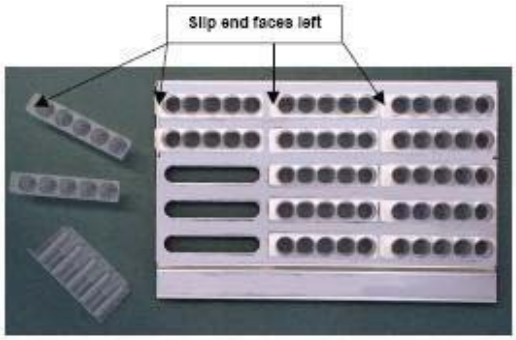
Step	Action
1	Place an Eppendorf tube per sample in the rack without the magnet inserted. Gently vortex the Dynabeads®anti O157, and add 20 µl Dynabeads®anti O157 to each tube. Use a lid opener for opening the lids of the Eppendorf tubes.
2	Gently add 1 ml of the pre-enriched sample to the Eppendorf tube. Use a new pipette / pipette tip for each sample. Close the lid. If sample is divided in 2 BPW bottles take 500 µl from each bottle.
3	Incubate the tubes for app. 10 minutes at room temperature. The rack is gently rotated without the magnet on a MX-3 Mixer (Dynal) or by hand.
4	Insert the magnet in the rack. Tilt the rack frequently for app. 3 minutes to ensure a complete collection of beads. With correct capture a distinct circular to oval brownish pellet is formed at the tube site halfway between the top and bottom of the tube.
5	Open the tubes gently by use of the lid opener. Place a Pasteur pipette at the water surface opposite to the pellet. Gently pipette up the supernatant and the liquid in the cap of the tube. Slow down pipetting when the surface of the liquid passes the pellet in order to make sure that no beads leave the tube through the pipette. If beads leave the sample, return the supernatant to the tube and repeat step 4. Use a new pipette / pipette tip for each sample.
6	Carefully add 1 ml of washing buffer to each sample. Do not touch the tube with the pipette / pipette tip since this can cross-contaminate the samples as well as the buffer. Close the lids and remove the magnet from the rack. Wash the bead complex by rotating the rack 3 times. Repeat step 4-6 twice, but the last time the pellet is only re-suspended in 100 µl washing buffer.

Continued on next page

Detection of E.coli in 25 g, *Continued*

Detection of E.coli O157 *(continued)*

Automatic ImmunoMagnetisk Separation (aIMS):

Step	Action
1	Load one sample tube for each sample into a sample rack.  *NOTE:* Each sample tube consists of 5 tubes called tube 1-5 (tube 1 is to the left (= slip end), and tube 5 is to the right).
2	Gently vortex the Dynabeads®anti O157 until the pellet in the bottom of the tube disappears, and aseptically add 10 µl properly mixed Dynabeads®anti O157 into sample tubes 1 and 2.
3	Aseptically add 500 µl of wash buffer to sample tubes 1 and 2. Aseptically add 1000 µl of wash buffer to sample tubes 3 and 4. Aseptically add 100 µl of wash buffer to sample tube 5.
4	Add 500 µl of the enriched test sample to sample tubes 1 and 2, be careful not to contaminate other tubes. If sample is divided in 2 BPW bottles take 500 µl from each bottle.
5	Repeat step 4 for the remaining samples.
6	Aseptically insert the sterile protective sample tip combs into the instrument.
7	Insert the rack with filled tubes into the instrument to lock it in place.
8	Check that everything is properly aligned. Close the instrument door
9	Select the EPEC/VTEC program sequence by scrolling with the arrow key, and press the Start button.

Continued on next page

Detection of E.coli in 25 g, *Continued*

Detection of E.coli O157 *(continued)*

Streaking onto selective indicative agar plates:

Each IMS product (from mIMS or from aIMS) is tested for the presence of *E.coli* O157 using selective indicative agar plates:

Step	Action
1	Gently vortex the pellet (IMS-product).
2	Streak 50 µl IMS-product onto the surface of a CT-SMAC agar plate, and streak another 50 µl IMS-product to the surface of a ChromAgar O157 plate (or a CT-ChromAgar O157 plate) in the following way: Spread the bead-bacteria complex over one half of the plate with a sterile cotton swab.This ensures the break-up of the bead-bacteria complexes. Dilute further by streaking with a loop.
3	Incubate the plates at 34-38°C for 18-24 hours.

Reading:

Agar	Description
CT-SMAC agar	On CT-SMAC agar, typical E.coli O157 colonies are transparent and almost colourless with a pale yellowish-brown appearance and a diameter of approximately 1 mm. Sorbitol positive organisms form bright red (pink) colonies. In some cases suspect colonies are so few that they can only be recognized in the bacterial lawn in the primary streaking zone. In this case, subculture suspect colony material onto a new CT-SMAC agar plate. If the growth is too weak after 18-24 hours, the plates can be re-incubated for up to 24 hours. In this case representative sorbitol negative colonies (transparent) shall be verified by use of the E.coli O157 Latex kit from Oxoid (see below).
ChromAgar O157 and CT-Chrom Agar O157	A typical *E.coli O157* will grow as a pink-mauve colony, whereas most other micro organisms are either inhibited or grow as blue or colourless colonies.

Continued on next page

Detection of E.coli in 25 g, *Continued*

Detection of E.coli O157 *(continued)*

Verification of *E.coli O157*:

Suspect colonies on CT-SMAC agar and ChromAgar O157 (or CT-ChromAgar O157) are verified as *E.coli O157* using O157 Latex test kit from Oxoid. The verification is performed according to the manufacturer's description.

NOTE: Dept. 402 may go directly to EV. E. coli PCR Verifcation (PSL-SM-3097)

Interpretation of results

E.coli Detected (DET) in 25 g

- Presence of typical colonies on TBX agar
- Presence of O157 Latex positive colonies from CT-SMAC agar and ChromAgar O157 (or CT-ChromAgar O157), i.e. suspect *E. coli* O157.

E.coli Not Detected (ND) in 25 g

- Absence of typical colonies on TBX agar
- Absence of O157 Latex positive colonies from CT-SMAC agar and ChromAgar O157 (or CT-ChromAgar O157), i.e. suspect *E. coli* O157.

Action on results

E.coli Detected (DET) in 25 g

- Report *E. coli* DET in 25 g
- Streak suspect colonies onto TSA agar and incubate 34-38°C for 18-24 h.
- Send TSA plate for verification to:

  Novozymes A/S Mikrobiologisk control Dept. 402
  Krogshoejvej 36, building 1KS.18
  DK-2880 Bagsvaerd
- Send an E-mail to "mkelab" and "JAah" (Responsible Scinetist) stating:

  E. coli Detected (subject line)
  LIMS no. of sample and ID no. of isolate
- Dept. 402 will enter verification result for EV. E. coli in LIMS directly for all sites if not otherwise agreed.
  -

E.coli Not Detected (ND) in 25 g

- Report *E. coli* ND in 25 g
- Report EV. *E. coli* ND in 25 g

NOTE: Kalundborg (DK) may send selective plate with suspect colonies.

Continued on next page

Detection of E.coli in 25 g, *Continued*

Sensitivity and specificity

Sensitivity: 100%
Specificity: 100%
REFERENCE: Luna doc.: 2006-04985 ; 2000-05132 ; 2007-28699

Filing

All documentation should be filed in accordance with the local filing SOP.

Contingencies

All deviations from this SOP should be discussed with the Method Responsible Scientist and should be documented.

References

ISO 16649-2 1st Ed. (2001) Microbiology of food and animal feeding stuffs - Horizontal method for the enumeration of presumptive *Escherichia coli* - Part 2: Colony-count technique at 44°C using 5-bromo-4-chloro-3-indolyl-β-D-glucoronic acid.

ISO 16654 1st Ed. (2001): Microbiology of food and animal feeding stuffs – Horizontal method for the detection of *Escherichia coli O157*.

The detection of *E.coli O157* is in accordance with ISO 16654 with the following exception:

- Enrichment is performed using Buffered Peptone Water at 34-38°C for 16-20 hours. ISO 16654 uses a modified TSB + novobiocin at 41.5°C ± 1°C for 18-24 hours.
- Immuno Magnetic Separation is only performed after 16-20 hours. ISO 16654 states after 6 hours and again, if necessary, after 12-18 hours (i.e. to a total elapsed time of 18-24 hours).
- Verification is performed using *E.coli O157* Latex test. ISO 16654 states indol test and serological test.

Revision

Major revision, this version supersede EB-SM-3007 and PSL-SM-3007.

Continued on next page

Detection of E.coli in 25 g, *Continued*

Flow chart

Flow chart of method, Click Link to read section.

Transfer 25 g sample into BPW bottle

Incubate BPW at 34-38°C for 16-20 h.

Plate on TBX agar plate

Incubate at 34-38°C for 18-24 h

Read TBX plate

Spread suspect colonies onto TSA

Incubate at 34-38°C for 18-24 h

Send plate for verification

Perform IMS (aIMS or mIMS)

Transfer 50 µl to CT-SMAC / CT Chrom

Spread with cottonswab and loop

Incubate at 34-38°C for 18-24 h

Read CT-SMAC and CT-Crom

Latex test on suspect colonies

Send plate with latex positive for verification

Result	Result Action
No suspect colonies on TBX **or** CT-SMAC/CT-Chrom	*E. coli* (25g) ND EV. *E. coli* ND
No suspect colonies on TBX **and** no latex positive from CT-SMAC/CT-Chrom	
Suspect colonies on TBX or latex positive colonies from CT-SMAC/CT-Chrom	*E. coli* (25g) DET Send plate for EV. *E. coli* verification

Detection of Salmonella spp.

Scope

All Novozymes QC laboratories involved in analysis of samples from Novozymes production and GLP studies.

Contents

Principle

Detection of *Salmonella* spp. is carried out as a qualitative test.

The test is based on a non-selective enrichment of 25 g sample in 450 ml Buffered Peptone Water for 1 day followed by selective enrichment using RVs broth or ImmunoMagnetic Separation (IMS, Manual mIMS or Automated aIMS)

The actual detection is carried out as an agar plate detection using XLD agar and Rambach agar. Suspect *Salmonella* colonies on the agar plates are verified using Oxidase test and API 20E or API Rapid 20E.

The methods are in-house methods evaluated and validated at Novozymes.

Definition of units

The result is stated as:

- DET (*Salmonella* detected in 25 gram)
- ND (*Salmonella* not detected in 25 gram)

Sample types

All Novozymes samples from production and GLP studies.

IMPORTANT: Hygiene samples are analyzed according to EB-SM-5001.

Continued on next page

Detection of Salmonella spp., *Continued*

Standards

A positive reference strain can be included in the test, e.g. *Salmonella adabraka*, *Salmonella havanna* or *Salmonella senftenberg* (Kalundborg isolates).

Detection limit

Theoretical detection limit: 1 *Salmonella spp.* in 25 g.

Safety

It is the responsibility of the laboratory leader that all personnel are aware of the correct handling of enzymes and reagents.

Equipment

Balance
Incubator for BPW and agar plates (34-38°C)
Incubator or water bath for RVs (40.0-42.0°C)
Vortex mixer
Pipettes and sterile tips (10-100 µl, 100-1000 µl, and 1 ml)
Sterile inoculation loops (1 and 10 µl)
For manual ImmunoMagnetic Separation (mIMS):
MPC-S Rack and magnet (Invitrogen Cat. No. 12020) +
Eppendorf tubes 1.5 ml (Eppendorf Cat. No. 0030 10086) +
Mixer, e.g. MX-3 Mixer (Invitrogen Cat. No. 15909). (Optional).
For automatic ImmunoMagnetic Separation (aIMS):
BeadRetriever (Invitrogen Cat. No. 15950) +
Tubes & tips (Invitrogen Cat. No. 15951)

Media and reagents

Buffered Peptone Water (BPW) (450 ml) prepared acc. to EB-ME-0009
Rappaport Vassiliadis soya peptone broth (RVs broth) (Oxoid CM0866)
IMS Dynabeads®anti-Salmonella (Invitrogen cat no.: 71002)
IMS Wash buffer (PBS-Tween 20 buffer) (Sigma No. P-3563)
XLD agar plates prepared acc. to EB-ME-0069
Rambach agar prepared acc. to EB-ME-0033
Tryptic Soy agar plates (TSA) prepared acc. to EB-ME-0041
Reagent for oxidase test, e.g. Bactident oxidase (Merck Cat. No. 1.13300.0001)
API Rapid 20E (BioMérieux Cat. No. 20 701) *or* API 20 E (BioMérieux Cat. No. 20 100) + relevant API reagents

Continued on next page

Detection of Salmonella spp., *Continued*

Non-selective enrichment

The non-selective enrichment is performed in the following way:

Step	Description
1	Transfer 25 g or 25 ml sample to 450 ml BPW equilibrated to room temperature. *IMPORTANT:* If any of these products transfer 25 g to 2 times 450 ml BPW bottles equilibrated to room temperature. 1. Novamyl SM30 conc. BG (AB…..) 2. Neutrase 1.5 Unstd MG (PW…..) 3. Ultraflo Unstd MG (CN…..) 4. Viscoflow MG (KR…..) 5. Flavourzyme 500 MG (HP…..) 6. Ceremix Plus MG (WD…..)
2	**If...** / **Then...** If Selective enrichment in RVs broth / Incubate BPW at 34-38°C for 16 - 24 hours If selective enrichment using IMS / Incubate BPW at 34-38°C for 24 ± 2 hours

Selective enrichment in RVs broth

The selective enrichment in RVs is performed in the following way:

- Transfer 100 µl *or* 0.1 ml from BPW to 10 ml RVs tubes equilibrated to minimum room temperature.
- Incubate the RVs broth at 40.0-42.0°C for 24 ± 2 hours.

NOTE: If 2 BPW bottles transfer 50 µl from each bottle to one RVs broth

NOTE: If water bath is used to incubate RVs there is no need to equilibrate the temperature of the RVs broth.

Continued on next page

Detection of Salmonella spp., *Continued*

Selective enrichment mIMS

The selective enrichment using mIMS is performed the following way:

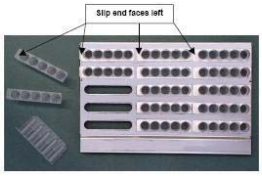
Step	Description
1	Place an Eppendorf tube per sample in the rack without the magnet inserted. Gently vortex the Dynabeads®anti-Salmonella and add 20 µl Dynabeads®anti-Salmonella to each tube.
2	Gently add 1 ml of the pre-enriched sample to the Eppendorf tube. Use a new pipette / pipette tip for each sample. Close the lid. If sample is enriched in 2 BPW bottles transfer 500 µl from each bottle.
3	Incubate the tubes for app. 10 min. at room temperature. The rack is gently rotated without the magnet on a MX-3 Mixer (or similar) or by hand.
4	Insert the magnet in the rack. Tilt the rack frequently for app. 3 min. to ensure a complete collection of beads. With correct capture a distinct circular to oval brownish pellet is formed at the tube site halfway between the top and bottom of the tube.
5	Open the tubes gently by use of the lid opener. Place a Pasteur pipette at the water surface opposite to the pellet. Gently pipette up the supernatant and the liquid in the cap of the tube. Slow down pipetting when the surface of the liquid passes the pellet in order to make sure that no beads leave the tube through the pipette. If beads leave the sample, return the supernatant to the tube and repeat step 4. Use a new pipette / pipette tip for each sample.
6	Carefully add 1 ml of washing buffer to each sample. Do not touch the tube with the pipette / pipette tip since this can cross-contaminate the samples as well as the buffer. Close the lids and remove the magnet from the rack. Wash the bead complex by rotating the rack 3 times. Repeat step 4-6 twice, but the last time the pellet is only re-suspended in 100 µl wash buffer.

Continued on next page

Detection of Salmonella spp., *Continued*

Selective enrichment aIMS

The selective enrichment using aIMS is performed the following way:

Step	Description
1	Load one sample tube for each sample into a sample rack.  Each sample tube consists of 5 tubes called tube 1-5 (tube 1 is to the left (= slip end), and tube 5 is to the right).
2	Gently vortex the Dynabeads®anti-Salmonella until the pellet in the bottom of the tube disappears, and aseptically add 10 µl properly mixed Dynabeads®anti-Salmonella into sample tubes 1 and 2.
3	Aseptically add 500 µl of wash buffer to sample tubes 1 and 2. Aseptically add 1000 µl of wash buffer to sample tubes 3 and 4. Aseptically add 100 µl of wash buffer to sample tube 5.
4	For each sample remove the labelled sample tube strip from the sample rack, and place it in a second sample rack. Add 500 µl of the enriched test sample to sample tubes 1 and 2, and return the inoculated tube strip to the first sample rack. If sample is enriched in 2 BPW bottles transfer 500 µl from each bottle. *CAUTION:* Be careful not to cross contaminate, if possible place racks well separated.
5	Repeat step 4 for the remaining samples.
6	Aseptically insert the sterile protective sample tip combs into the instrument.
7	Insert the rack with filled tubes into the instrument to lock it in place.
8	Check that everything is properly aligned. Close the instrument door
9	Select the Salmonella program sequence by scrolling with the arrow key, and press the Start button.

NOTE: Check that all magnets have been transferred to tube 5 as this may not occur always with some difficult sample types. If all magnets have not been transferred to tube 5, then repeat step 9. If the problem remains then repeat step 1-9 using four sample tubes instead of one. Add only 125 µl enriched test sample to tubes 1 and 2 in each sample tube. Transfer all material from tubes no. 5 into one of the no. 5 tubes (approx. 400 µl in total).

Continued on next page

Detection of Salmonella spp., *Continued*

Detection

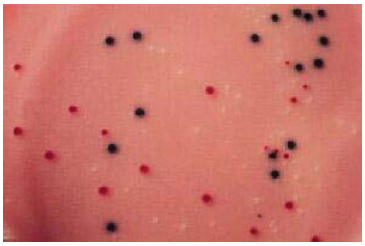
RVs broth is tested for the presence of *Salmonella spp.* using two different selective indicative agar plates (XLD agar and Rambach agar):

<table>
<tr><th>Step</th><th>Description</th></tr>
<tr><td>1</td><td>
<table>
<tr><th>If...</th><th>Then...</th></tr>
<tr><td>RVs</td><td>Mix (vortex) RVs broth.</td></tr>
<tr><td>IMS</td><td>Mix (vortex) the pellet IMS-product.</td></tr>
</table>
</td></tr>
<tr><td>2</td><td>
<table>
<tr><th>If...</th><th>Then...</th></tr>
<tr><td>RVs</td><td>Streak 10 µl RVs broth using a 10 µl inoculation loop onto the surface of a XLD agar plate, and streak another 10 µl RVs broth to the surface of a Rambach agar plate. Same inoculation loop may be used.</td></tr>
<tr><td>IMS</td><td>Streak 50 µl IMS-product onto the surface of a XLD agar plate, and streak another 50 µl IMS-product to the surface of a Rambach agar plate.
Spread the bead-bacteria complex over one half of the plate with a sterile cotton swab. This ensures the break-up of the bead-bacteria complexes. Dilute further by streaking with an inoculation loop (1 µl).
NOTE: If IMS has been performed using four sample tubes then streak 200 µl IMS-product on each of the XLD and Rambach agar plates.</td></tr>
</table>
</td></tr>
<tr><td>3</td><td>Incubate the plates at 34-38°C for 1 day (minimum 18 hours).</td></tr>
</table>

Continued on next page

Detection of Salmonella spp., *Continued*

Reading plates

Agar	Description
Rambach agar	 : Na-desoxycholate inhibit gram-positive flora. Indicative principle: Contains pH indicator, propylene glycol and chromogene. Rambach agar Salmonella • Red colonies – produce acid from propylene glycol. *NOTE:* *S. arizona* form brownish, green-purple or blue-purple colonies. Coliform • Blue-violet/Blue-green colonies – presence of β-D-galactosidase. Other enterobacteriaceae and gram-negative bacteria (Proteus, Pseudo-monas, Shigella, S. typhi and S. paratyphi A) • Colourless-yellow colonies.
XLD	 • : Na-desoxycholate inhibits gram-positive flora and coliforms. • Indicative principle: Contains lysine and H_2S indicators. • • XLD agar Salmonella • Red colonies with a black centre (often seen as black colonies) – black colour due to H_2S-production. Some strains of *S. havana* form gray-brownish colonies (“fish-eye”) as they are H_2S-negative. *NOTE:* Be aware of very small black pin-point Salmonella colonies. Shigella, Providencia, H_2S-negative Salmonella (some Proteus and Pseu-domonas) • Red colonies.

NOTE: Rambach and XLD agar may be stored at cool for up to 48 hours before reading cf. (Ref. 2).

Continued on next page

Detection of Salmonella spp., *Continued*

Verification

IMPORTANT: Verification using **API Rapid 20E** tests must always be performed using colonies subcultivated on TSA agar plates incubated at 34-38°C for 1 day.

IMPORTANT: Verification using **API 20E** is traditionally executed with pure colonies cultivated on a non selective agar plate such as TSA agar. However, according to (Ref. 2) it is possible to perform verification using API 20E directly from the
selective XLD and Rambach agar plates.

Hence, verification using API 20E is performed as follows:

<table>
<tr><th>Step</th><th colspan="2">Description</th></tr>
<tr><td>Day 1</td><td>If...</td><td>Then...</td></tr>
<tr><td>A</td><td>A suspect colony is present as a single, pure, colony on Rambach or XLD</td><td>1. Perform API 20E test using a single, pure colony, from Rambach or XLD agar plate.
2. Subcultivate from the same colony on a TSA agar plate. Incubate at 34-38°C for 1 day.</td></tr>
<tr><td>B</td><td>A suspect colony is present but <u>not</u> as a single, pure, colony on Rambach or XLD</td><td>1. Streak suspect colony onto new Rambach and/or XLD agar plate and a TSA agar plate. Incubate all plates at 34-38°C for 1 day.
2.
<table>
<tr><th>If...</th><th>Then...</th></tr>
<tr><td>Pure colony on TSA[1]</td><td>Proceed with step Day 2B.</td></tr>
<tr><td>Pure colony on Rambach or XLD</td><td>Proceed with step Day 1A.</td></tr>
</table>
[1] Caution: It may be difficult to identify a Salmonella colony on TSA agar if more than one colony type is present on the agar plate.</td></tr>
</table>

Continued on next page

Detection of Salmonella spp., *Continued*

Verification *(continued)*

<table>
<tr><th>Step</th><th colspan="3">Description</th></tr>
<tr><td>Day 2</td><td>If...</td><td colspan="2">Then...</td></tr>
<tr><td>A</td><td>A suspect colony is present as a single, pure, colony on Rambach or XLD</td><td colspan="2">1. Perform oxidase test on colony material from TSA agar plate.
2.<table><tr><th>If...</th><th>Then...</th></tr><tr><td>Oxidase test is positive</td><td>Result is ND</td></tr><tr><td>Oxidase test is negative</td><td>Read API 20E test and determine API ID.</td></tr></table></td></tr>
<tr><td>B</td><td>A suspect colony is present but not as a single, pure, colony on Rambach or XLD</td><td colspan="2">1. Perform oxidase test on colony material from TSA agar plate.
2.<table><tr><th>If...</th><th>Then...</th></tr><tr><td>Oxidase test is positive</td><td>Result is ND</td></tr><tr><td>Oxidase test is negative</td><td>Perform API 20 E test using a single, pure colony, from TSA agar plate.</td></tr></table></td></tr>
<tr><td>Day 3</td><td>If...</td><td colspan="2">Then...</td></tr>
<tr><td>B</td><td>A suspect colony is present but not as a single, pure, colony on Rambach or XLD</td><td colspan="2">Read API 20E and determine API ID.</td></tr>
</table>

IMPORTANT: Verification must always be performed using freshly grown cultures, i.e. verification may not be performed from agar plates stored at cool. If selective agar plates have been stored at cool before reading, fresh cultures must be prepared by subcultivation of suspect colonies on a TSA agar plate incubated at 34-38°C for 1 day.

Continued on next page

Detection of Salmonella spp., *Continued*

Verification *(continued)*

Verification Kits

Local procedure SOP for Oxidase and API may be used, e.g. Oxidase test (DK) and API 20E (DK).

Oxidase test (e.g. Bactident Oxidase, Merck Cat. No. 1.13300.0001)

a. Remove a single isolated, well-developed colony from the culture medium with a loop.
b. Apply the colony to the reactive zone of the oxidase strip and distribute with the aid of the loop.
c. After 20-60 seconds compare the test strip with the colour scale provided. If cytochrome c oxidase-positive bacteria are present the reactive zone exhibits a blue to purple colour. If cytochrome c oxidase-negative bacteria are present the reactive zone exhibits remains colourless.

NOTE: If using comparable test, please follow manufacturer directions.

Continued on next page

Detection of Salmonella spp., *Continued*

Verification *(continued)*

API Rapid 20E (a 5 hour test)

a. Transfer 1-4 colonies to a “API NaCl 0.85%, 2 ml” vial (corresponding to McFarland 0.5), and mix carefully.

b. Inoculate the API Rapid 20 E strip: With the same pipette, distribute the suspension into the tubes of the strip. To avoid the formation of bubbles at the base of the tube, tilt the strip slightly forwards and place the tip against the side of the cupule.
- For the CIT test, add 2 drops of the suspension (app. 50 µl) to fill the tube and lower position of the cupule.
- For the other tests, only fill the tubes (app. 50 µl per tube). The accuracy of the filling is very important.
- For the underlined tests (LDC, ODC and URE) completely fill the cupule with mineral oil.

c. Incubate the strip at 34-38°C for 4-4½ hours.

d. Read the strips by referring to the reading table (in the package insert) and the picture below:
- VP test (performed in a safety bench wearing protective gloves): add 1 drop of each of VP 1 and VP 2 reagents. Wait 5-10 minutes. A red color indicates a positive reaction.
- IND test: add 1 drop of JAMES reagent. The reaction takes place immediately. A red colour indicates a positive reaction.

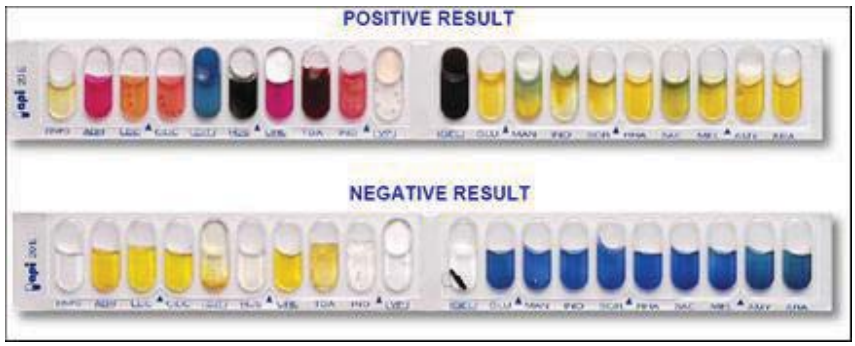
Continued on next page

Detection of Salmonella spp., *Continued*

Verification *(continued)*

API 20E (a 2 days test)

a. Transfer 1 colony to a “API NaCl 0.85%, 5 ml” vial, and mix carefully.
b. Inoculate the API 20 E strip: With the same pipette, distribute the suspension into the tubes of the strip. To avoid the formation of bubbles at the base of the tube, tilt the strip slightly forwards and place the tip against the side of the cupule.
   - For the CIT, VP and GEL tests, add 2 drops of the suspension (app. 50 µl) to fill the tube and lower position of the cupule.
   - For the other tests, only fill the tubes (app. 50 µl per tube). The accuracy of the filling is very important.
   - For the underlined tests (ADH, LDC, ODC, H_2S and URE) completely fill the cupule with mineral oil.
c. Incubate the strip at 34-38°C for 18-24 hours.
d. Read the strips by referring to the reading table (in the package insert) and the picture below:
   - TDA test: add 1 drop of TDA reagent. A red / brown color indicates a positive reaction.
   - IND test: add 1 drop of JAMES reagent. The reaction takes place immediately. A red color indicates a positive reaction.
   - VP test (performed in a safety bench wearing protective gloves): add 1 drop of each of VP 1 and VP 2 reagents. Wait 5-10 minutes. A red colour indicates a positive reaction.



Continued on next page

Detection of Salmonella spp., *Continued*

API Identification

Read and determine ID using API webb:

Step	Action
1	Take out an API worksheet corresponding to the appropriate API strip (20E or Rapid 20E) and mark all the positive and negative results with a + or - .
2	Calculate the total score for each section of three tests on the API worksheet, only positives are tabulated. This will result in a 7 digit profile number.
3	Go to the website below, log-in, and choose the correct API test (either 20E or Rapid 20E): apiwebb™
4	Enter the number for the appropriate group of 3 and hit confirm. Your Identification will appear. *NOTE:* For API 20E, a correct identification require ≥ 80% similiarity. Contact responsible chemist if the ID score is < 80%.

IMPORTANT: In the LDC (Lysine decarboxylase) reaction it can be difficult to distinguish between Yellow/Negative and Orange/Positive. *Citrobacter braakii* may be wrongly identified as Salmonella based on this test. *C. Braakii* is Negative for LDC and Salmonella is Positive. If API identification is Salmonella Detected and LDC is Negative re-test and contact Responsible Scientist.

Interpretation of results

If…	Then report result as…
No suspect colonies on Rambach agar and XLD agar	*Salmonella* spp. not detected (ND)
Suspect colony on Rambach agar or XLD and colony is oxidase positive	*Salmonella* spp. not detected (ND)
Suspect colonies on Rambach agar or XLD and colony is oxidase negative but API Webb ID score is < 80%	*Salmonella* spp. not detected (ND)
Suspect colonies on Rambach agar or XLD and colony is oxidase negative and API Webb ID score is ≥ 80%	*Salmonella* spp. detected (DET)

Sensitivity and specificity

Sensitivity: 100% Specificity: 100%

REFERENCE: Luna No. 2008-20805-01.

Continued on next page

Detection of Salmonella spp., *Continued*

Filing

All documentation should be filed in accordance with the local archiving SOP.

Contingencies

All deviations from this SOP should be discussed with the Method Responsible Scientist and should be documented.

References

1. LUNA No. 2008-20805-01: Development and validation of two new methods for detection of *Salmonella spp.* in enzyme samples.
2. LUNA No. 2012-02028-01: Hurtigere påvisning af Salmonella samt aflæsning af selective plader opbevaret på køl. *In Danish*t

Rambach agar, Merck
XLD agar, Oxoid

Revision

The section “Materials” has been subdivided into the sections “Equipment” and “Media and reagents”. The product “Clear Lens Pro 2.5 MG“ has been deleted from the section “non-selective enrichment” as it is no longer produced in NZ. Added notes on analysis of troublesome sample types in the sections “Selective enrichment aIMS” and “Detection”. Added that RVs broth may be vortexed in the section “Detection”. Added that Salmonella arizona form brownish, green-purple or blue-purple colonies and that selective agar plates may be stored at cool for up to 48 hours before reading in the Section “Reading plates”. Major revision of the section “Verification” adding the option of performing verification on colonies from selective agar plates cf. Ref. 2 . Added oxidase test result of cytochrome c oxidase-negative bacteria. The section “Interpretation of results” has been rewritten to include interpretation of verification analyses. Added reference to trainee report in the section “References”. Revised flow charts.

Continued on next page

Detection of Salmonella spp., *Continued*

Flow Chart RVs

Flow chart of method using RVs broth, Click Ctrl + to read section.

Transfer 25 g/ml sample into 450 ml BPW bottle
Incubate BPW at 34-38°C for 16 - 24 hours.

Transfer 100 µl or 0.1 ml to RVs tube
Incubate the RVs at 40.0-42.0°C for 24 ± 2 hours.

Streak RVs with 10 µl inoculation loop onto XLD and Rambach
Incubate plates at 34-38°C for 1 day (minimum 18 hours)

No suspect colonies on either XLD or Rambach = **Salmonella ND**

Suspect colonies on either XLD or Rambach

Perform verification

Interpretate results

Pos. ox. test = **Salmonella ND**
Neg. ox. test + API ID < 80% = **Salmonella ND**
Neg. ox. test + API ID ≥ 80% = **Salmonella DET**

Continued on next page

Detection of Salmonella spp., *Continued*

Flow Chart IMS

Flow chart of method using IMS, Click Ctrl + to read section.

Transfer 25 g/ml sample into 450 ml BPW bottle.
Incubate BPW at 34-38°C for 24 ± 2 hours.

Perform IMS (aIMS or mIMS)

Streak 50 µl IMS-product on both an XLD agar and Rambac agar plate

Incubate plates at 34-38°C for 1 day (minimum 18 hours)

No suspect colonies on either XLD or Rambach = **Salmonella ND**

Suspect colonies on either XLD or Rambach

Perform verification

Interpretate results

Pos. ox. test = **Salmonella ND**
Neg. ox. test + API ID < 80% = **Salmonella ND**
Neg. ox. test + API ID ≥ 80% = **Salmonella DET**

Detection of Antimicrobial activity

Scope

All Novozymes QC laboratories involved in analysis of samples from Novozymes production and GLP studies.

Principle

Detection of Antimicrobial activity is based on the measurement of inhibition of bacterial growth under specific circumstances.

The method is in accordance with JECFA (1992)

LIMS code

ANTIMIC

Definition of units

The result is stated as:

- DET (Antimicrobial activity detected) *or*
- ND (Antimicrobial activity not detected)

Samples

All sample types.

Standards

Staphylococcus aureus, ATCC 6538
Escherichia coli, ATCC 11229
Bacillus cereus, ATCC 2
Bacillus circulans, ATCC 4516
Streptococcus pyogenes, ATCC 12344
Serratia marcescens, ATCC 14041
NOTE: The test organisms must be traceable.

Detection limit

Not known.

Continued on next page

Detection of Antimicrobial activity, *Continued*

Equipment

Balance (± 0.1 g)
Sterile pipettes for transfer of 100 µl, 1 ml and 10 ml
Inoculation loops 1µl
Paper discs, e.g. S&S Analytical Filter Papers No. 740-E (12.7 mm in diameter), autoclaved
Bio Safety Cabinet, Class II
Sterile gloves
Refrigerator (2-8°C)
Incubator (34-38°C)
-80°C freezer
Ruler or Vernier gauge
Petri dishes, 9 cm

Media and reagents

Tween buffer 4%
Tryptone Soya agar (TSA), 90 ml in 250 ml Blue cap bottles
Tryptone Soya agar plates, 9 cm with app. 15 ml agar (TSA)
CASO broth, 50 ml
IMPORTANT: Preparation in the local laboratory shall be done according to the current valid WW Media direction.
Ciprofloxacin discs (5 µg or 10 µg) (bought ready to use).

Safety

It is the responsibility of the laboratory leader that all personnel are aware of the correct handling of enzymes and reagents.

Continued on next page

Detection of Antimicrobial activity, *Continued*

Day 1: Preparation of test organisms in CASO

Handling the test organisms must be performed in a Bio Safety Cabinet, Class II.

Step	Action
1	Inoculate each of the 6 test organisms using a 1 µl inoculation loop (the strains are taken directly from a Cryo tube that has been stored in a -80°C freezer) in separate CASO broth, 50 ml.
2	Contemporary, streak out each test organism, using the same inoculation loop as in step 1, on the surface of a TSA plate to look for purity.
3	Incubate the CASO broth and TSA plates overnight at 34-38°C.

Day 2: Number of test organisms in CASO

The number of test organisms is tested in each of the CASO broths.

Step	Action
1	Make a 10^{-4} dilutions of the following test organisms: • Bacillus cereus • Bacillus cirkulans
2	Make a 10^{-5} dilutions of the following test organisms: • Staphylococcus aureus • Escherichia coli • Streptococcus pyrogenes • Serratia marcescens
3	Determine the Total viable count of each dilution by spread plate or spiral plate on TSA plates. Incubate overnight at 34-38°C

Continued on next page

Detection of Antimicrobial activity, *Continued*

Day 2: Infection control (purity of the Cryo tubes)

Purity of the test organisms from the Cryo tubes are controlled the following way:

Step	Action
1	Control the purity of the 6 test organisms on TSA plates from the day before.
2	Write down the result (+ or – infection).

Day 2: Preparation of test plates and purity of the test organisms in CASO broth

Preparation of test plates must be done in a Bio Safety Cabinet, Class II and wearing sterile gloves.

Step	Action
1	For each test organism a bottle containing 90 ml of Tryptone Soya agar (TSA) is melted and cooled (to app. 47°C)
2	Transfer 10 ml of CASO broth inoculated with *S. aureus* to a 250 ml Blue cap bottle with 90 ml melted and cooled Tryptone Soya agar (TSA). Mix carefully.
3	Pour app. 10 ml of the TSA-microorganism mixture onto an already prepared TSA plates (containing app. 15 ml TSA). Distribute the TSA-microorganism mixture evenly on the surface of the TSA plate, and allow solidifying.
4	Make another 9 plates as described in step 3.
5	Control the purity of the CASO broth by streaking out form the last drop of the bottle with a 1 µl inoculation loop onto the surface of one TSA plate.
6	Repeat step 2-5 for the rest of the microorganisms. *IMPORTANT:* Transfer only 5 ml of the CASO broth containing *Streptococcus pyogenes* to 90 ml of melted and cooled TSA.
7	Incubate the TSA plates prepared in step 5 overnight at 34-38°C.

Continued on next page

Detection of Antimicrobial activity, *Continued*

Day 2: Control of test plates with Ciprofloxacin

Control the test plates with Ciprofloxacin to determine whether the test organisms are capable of making an inhibition zone on the test plate.

The control must be done on one test plate per test organism.

Step	Action
1	Put one disc of Ciprofloxacin onto the middle of a test plate.
2	Place the test plate overnight at 2-8°C.
3	Incubate the test plate overnight at 34-38°C.

Day 3: Purity in CASO broth

Check the purity of the TSA plates from the day before.

Write down the result of the purity test for each of the test organisms (+ or – infection).

Day 3: Number of colonies on TSA plates

Count the number of colonies on the TSA plates from the day before.

Write down the results.

IMPORTANT: To approve the test plates all readings must be $>10^6$ CFU.

Day 3: Reading of inhibition zone

Reading of the TSA plates with Ciprofloxacin from the day before is done by measuring the diameter of the inhibition zone on each of the test plates using a ruler or a Vernier gauge.

Write down the results.

IMPORTANT: Each zone must be ≥25 mm.

Sample preparation

The samples are prepared as followed:

- Transfer 10 g of solid sample or 10 ml of liquid sample to 90 ml Tween buffer 4%
- Immediately homogenize the sample by stirring or by shaking. Solid samples are homogenized on a magnetic stirrer for app. 20 minutes

Continued on next page

Detection of Antimicrobial activity, *Continued*

Test procedure

The test is performed in the following way:

TIP: It is recommended to work with dry plates, and to place the disc on the test plate immediately before addition of sample so the filters do not absorb moisture from the plates and thus cannot absorb the sample.

Step	Action
1	Place a sterile paper disc on each of the 6 test plates (one test plate per micro organism). *NOTE:* Up till 5 sterile paper discs can be placed on one plate (giving the possibility of analysing up to 5 samples per set of 6 test plates).
2	Inoculate each paper disc with 100 µl of the 10^{-1} dilution of the sample prepared above.
3	Place the plates overnight at 2-8°C.
4	Incubate the plates overnight at 34-38°C.
5	Measure the diameter of the inhibition zone on each of the plates using a ruler or a Vernier gauge.
6	Write down the results (inhibition zone in mm.).

Continued on next page

Detection of Antimicrobial activity, *Continued*

Interpretation of results

Results are given according to the sheet below:

Is there…	with a zone measuring…	…the result is
0 inhibition zones	0 mm	Not detected (ND)
X inhibition zones	<16 mm	Not detected (ND)
1 inhibition zones	≥16 mm	Not detected (ND)
2 inhibition zones	≥16 mm	Not detected (ND)
3 inhibition zones	≥16 mm	Detected (DET)

If the result is Detected (DET) a remark is given on which of the test organisms that shows obvious antimicrobial activity in the sample and the size of the zone is stated.

IMPORTANT: If the result is Detected (DET) the Responsible Scientist is contacted.

Accuracy and precision

Not validated.

Archiving

All documentation should be archived in accordance with the local archiving SOP.

Contingencies

All deviations from this SOP should be discussed with the Method Responsible Scientist and should be documented.

References

Joint FAO/WHO Expert Committee on Food Additives (JECFA). Compendium of food additive specifications, Volume 1, Rome 1992, appendix A to annex 1.

Revision

Both 5 µg and 10 µg Ciprofloxacin discs can be used (Luna no. 2008-31511)

Mycotoxins by LC-MS/MS

Purpose

This document details the analysis of multiple mycotoxins by Liquid Chromatography coupled to triple quadrupole mass spectrometer (LC/MS/MS). The method covers enzymes concentrates, enzymes fermentations and wheat bran.

Scope

PSL (Process Support Laboratories; 575) and PAS (Pharma Analytical Science; 1000-367)

Appendix

In the back of this document please find an Appendix with hyperlinks to external documents relevant for this method as well as supplementary information.

Overview

Overview of bookmarks in this document:

Principle

The sample is precipitated and mycotoxins are extracted with formic acid and acetonitrile. The extract is centrifuged and the supernatant is analysed on LC-MS/MS.

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Safety and hazards

Latex-free gloves are used. All work with toxins should be performed in fume-hood. Toxins are marked from corrosive (C) to very toxic (Tx). There are components which can have carcinogenic effect. Therefore stocksolutions and pure substances are handled with great caution. Use plastic apron when preparing stock solution. Read APB as well as other safety documents before starting work. See Appendix. For general working environment reference to local procedure.

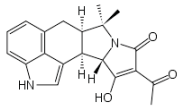
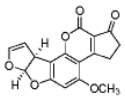
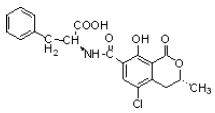
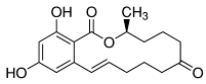
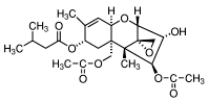
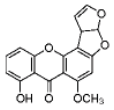
Sample types

This method is suitable for: enzyme concentrates, enzyme fermentations, Wheat Bran

Chemicals

Other chemicals than the ones listed below can be applied if the quality is equivalent, but chromatography ought to be tested prior to changing product. See Appendix for specification of which chemicals are used.

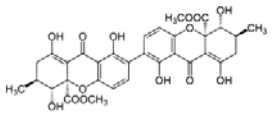
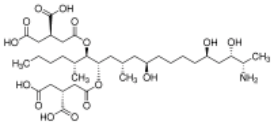
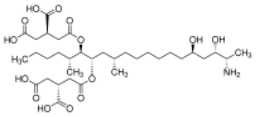
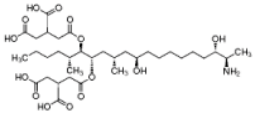
Chemicals used for stock and standards:

Chemical	Abbreviation	Structure
Cyclopiazonic acid	**CPA**	
Aflatoxin B1	**AFB1**	
Ochratoxin A	**OTA**	
Zearalenone	**ZEA**	
T2 toxin	**T2**	
Sterigmatocystin	**STE**	

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Chemicals *(continued)*

Chemical	Abbreviation	Structure
Secalonic acid D	**SAD**	
Fumonisin B1	**FB1**	
Fumonisin B2	**FB2**	
Fumonisin B3	**FB3**	

CoA of the chemicals should be available when possible.

Chemicals used for eluents and mobile phases:

Chemical	Abbreviation
Acetonitrile	ACN
Formic Acid	HCOOH
Methanol	MeOH

Reagents

Reagent	Preparation
Eluent A: (A1) ACN: H_2O: HCOOH (10:90:0.15)	Appendix
Eluent C: (B1) ACN: HCOOH (100:0.15)	

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Stock solutions

Information about preparation is documented, see Appendix for hyperlink.

Prepare two stock solutions based on different weighings of material; one solution for standards and one solution for QC samples.

Cyclopiazonic acid: 50 ppm

Step	Action
1	Weigh 10 ± 1 mg CPA and transfer quantitatively to a 200 mL volumetric flask. (If chemical is available in ampoule with certified amount, weighing is not required)
2	Add **methanol** to 200 mL and mix

Aflatoxin B_1: 50 ppm

Step	Action
1	Transfer 5 mg AFB1 quantitatively to 100 mL volumetric flask. *NOTE:* This is done by dissolving the substance directly in the ampoule with approx. 3 x 2 mL acetonitrile.
2	Add **methanol** to 100 mL and mix.
3	Store solution **dark** (Aflatoxin degradates in light).

Ochratoxin A: 50 ppm

Step	Action
1	Transfer 5 mg OTA quantitatively to 100 mL volumetric flask. *NOTE:* This is done by dissolving the substance directly in the ampoule with approx. 3x2 mL **methanol**
2	Add **methanol** to 100 mL and mix

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Stock solutions *(continued)*

Zearalenon: 50 ppm

Step	Action
1	Transfer 5 mg ZEA quantitatively to 100 mL volumetric flask. *NOTE:* This is done by dissolving the substance directly in the ampoule with approx. 3x2 mL methanol
2	Add **methanol** to 100 mL and mix

T-2 toxin: 50 ppm

Step	Action
1	Transfer 5 mg T2 quantitatively to 100 mL volumetric flask. *NOTE:* This is done by dissolving the substance directly in the ampoule with approx. 3x2 mL **methanol**
2	Add **methanol** to 100 mL and mix

Sterigmatocystin: 50 ppm

Step	Action
1	Transfer 5 mg STE quantitatively to 100 mL volumetric flask. *NOTE:* This is done by dissolving the substance directly in the ampoule with approx. 3 x 2 mL **acetonitrile**.
2	Add **methanol** to 100 mL and mix.

Secalonic Acid D: 50 ppm

Step	Action
1	Weigh 10 ± 1 mg SAD and transfer quantitatively to a 200 mL volumetric flask with 2-3 ml **acetone**.
2	Add **methanol** up to 200 mL og mix

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Stock solutions *(continued)*

Fumonisin B1, B2, B3: 50ppm

Step	Action
1	Fumonisin stock solutions 50µg/ml are used directly according to the table below.

IMPORTANT: Portion each component into small bottles. Dispose of volumetric flask and cap. Mark bottles with component name, R-number, date, expiry date, toxic, highly inflammable**.**

IMPORTANT: **The QC-sample should be prepared from a different weighing of chemicals.**

Standard Mix Stock

The Standard Mix Stock solution is prepared as followes:

Step	Action		
1	Pipette, into a 50 mL volumetric flask, following volumes of stock solutions.		
	CPA	1500 µL	1.5 ppm
	Aflatoxin B_1	200 µL	0.2 ppm
	Ochratoxin A	200 µL	0.2 ppm
	Zearalenon	400 µL	0.4 ppm
	T-2 toxin	400 µL	0.4 ppm
	Sterigmatocystin	200 µL	0.2 ppm
	Secalonic Acid D	1500 µL	1.5 ppm
	Fumonisin B1	200 µL	0.2 ppm
	Fumonisin B2	200 µL	0.2 ppm
	Fumonisin B3	200 µL	0.2 ppm
2	Add methanol to 50 mL and mix.		
3	The solution is marked toxic and highly inflammable.		
4	Document the preparation of stock solution in a validated spreadsheet found in Appendix		

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

MeOH standards

MeOH standards 1-7

Step	Action
1	Prepare following standards according to the table below, using Mix Stock

MeOH std	Stock (µL)	MeOH to (mL)	Conc. *) (µg/L)		
			CPA + SAD	T2 + ZEA	AFB1 + STE + OTA + FUM1 + FUM2 + FUM3
1	35	25	2.1	0.56	0.28
2	70	25	4.2	1.1	0.56
3	140	25	8.4	2.2	1.1
4	350	25	21	5.6	2.8
5	700	25	42	11	5.6
6	1400	25	84	22	11
7	2800	25	168	45	22

2	Document the preparation in a validated spreadsheet found in Appendix
3	Transfer 500 µl methanol std 1-7 to vials and mark them MeOH STD 1-7
4	The box with vials is placed in freezer and is marked with preparation date, expiring date, toxic and highly inflammable.

*) The accurate concentration of CPA, Secalonic Acid and Fumonisin B1, B2, B3 is being calculated based on weighing numbers and is transferred to validated spreadsheet found in Appendix.

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Standards

Working standard 1 – 7:

Step	Action
1	Prepare following standards according to table below, by adding 500 µl MilliQ water to MeOH STD 1-7 vials. This is done prior to analysis.
	See table below
2	Document the preparation on data capture template (link found in Appendix)
3	Analyse on LC-MS/MS

STD	MeOH STD	Vol. water (µL)	Conc. *) (µg/L)		
			CPA + SAD	T2 + ZEA	AFB1 + STE + OTA + FUM1 + FUM2 + FUM3
1	1	500 µL	1.1	0.28	0.14
2	2	500 µL	2.1	0.56	0.28
3	3	500 µL	4.2	1.1	0.56
4	4	500 µL	11	2.8	1.4
5	5	500 µL	21	5.6	2.8
6	6	500 µL	42	11	5.6
7	7	500 µL	84	22	11

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

QC Mix stock

The QC Mix Stock solution is also used as Spike solution. The solution is prepared as follows:

Step	Action
1	Pipette, into a 50 mL volumetric flask, following volumes of stock solutions.
	CPA 15000 µL 15 ppm Aflatoxin B_1 200 µL 0.2 ppm Ochratoxin A 200 µL 0.2 ppm Zearalenon 400 µL 0.4 ppm T-2 toxin 400 µL 0.4 ppm Sterigmatocystin 200 µL 0.2 ppm Secalonic Acid D 15000 µL 15 ppm Fumonisin B1 200 µL 0.2 ppm Fumonisin B2 200 µL 0.2 ppm Fumonisin B3 200 µL 0.2 ppm
2	Add methanol to 50 mL and mix.
3	The solution is marked toxic and highly inflammable.
4	Document the preparation of stock solution in a validated spreadsheet found in Appendix

NOTE: If possible the QC Mix Stock should origin from stock solutions different from the stock solutions used for the preparation of Standard Mix Stock solution.

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

MeOH QC-samples

MeOH QC samples are prepared as follows:

Step	Action
1	Prepare following standards according to the table below, using QC Mix Stock

MeOH std	Stock	MeOH	Conc. *)

		(µL)	**to (mL)**	**(µg/L)**		
				CPA + SAD	**T2 + ZEA**	**AFB1 + STE + OTA + FUM1 + FUM2 + FUM3**
	QC	140	25	120	3.2	1.6
2	Document the preparation in a validated spreadsheet found in Appendix					
3	Transfer 500 µl methanol QC samples MeOH QC					
4	The box with vials is placed in freezer and is marked with preparation date, expiring date, toxic and highly inflammable.					

*) The accurate concentration of CPA, Secalonic Acid and Fumonisin B1, B2, B3 is being calculated based on weighing numbers and is transferred to validated spreadsheet found in Appendix.

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

QC-samples

QC-samples are prepared as follows:

Step	Action
1	Prepare following standards according to table below, by adding 500 µl MilliQ water to MeOH QCvials. This is done prior to analysis.
	<table><tr><th rowspan="2">STD</th><th rowspan="2">MeOH STD</th><th rowspan="2">Vol. water (µL)</th><th colspan="3">Conc. *) (µg/L)</th></tr><tr><th>CPA + SAD</th><th>T2 + ZEA</th><th>AFB1 + STE + OTA + FUM1 + FUM2 + FUM3</th></tr><tr><td>QC</td><td>QC</td><td>500 µL</td><td>60</td><td>1.6</td><td>0.80</td></tr></table>
2	Document the preparation on data capture template (link found in Appendix)
3	Analyse on LC-MS/MS

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Procedure

Step	Action
1	Document sample preparation in data capture template See Appendix for local hyperlink to template
2	Prepare four 10-ml Nunc tubes pr sample
3	**Solid samples:** Weigh 1.00 g (± 0.10 g) of the sample in each of the four 10 ml Nunc tubes **Liquid samples:** Pipette 1.00 mL of the sample in each of the four 10 ml Nunc tubes
4	Note weight or volume
5	Provide the four sub-samples with the suffixes “a”, “b”, “as” and “bs”
6	Spike “as” and “bs” with 25 µl spike solution each
7	Add 200 µl formic acid to all samples
8	Add 1000 µl MQ-water to solid samples (wheat bran)
9	Add 3000 µl ACN to all samples
10	For solid samples: vortex vigorously for 5 seconds – ensure contact between solvents and samples
11	Shake samples for 30 minutes
12	Centrifuge (≥ 5 min, ≥ 3500 rpm)
13	Transfer 500 µl supernantant to a new vial and add 500 µl MilliQ water
14	Analyse on LC-MS/MS
15	If the peak area of a mycotoxin in a sample is above peak area of STD7, then dilute the sample in an appropriate volume of MQ and re-peat the analysis from *step 1*.

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Recommended dilution

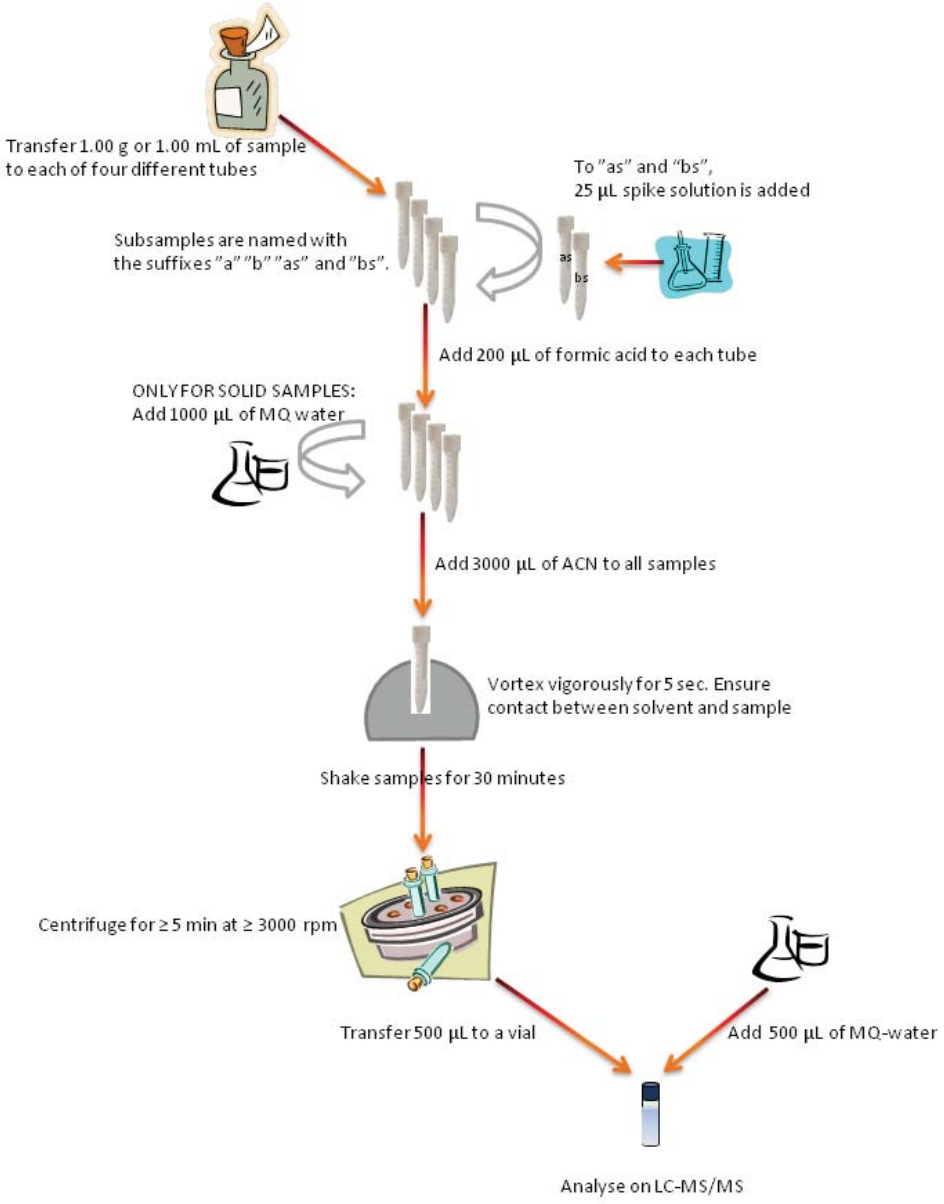
For some sample types it is necessary to perform a dilution prior to sample preparation when analyzing for CPA due to high content of this.

The recommended dilutions are as follows:

BATCH	DILUTION
LAR	1 in 2.5
LAD	1 in 5
CFD	1 in 500
CFR	1 in 50
CFG	1 in 50
CFZ	1 in 50
UAD	No dilution necessary
KEG	1 in 10

Mycotoxins by LC-MS/MS, *Continued*

Flow chart



Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Data collection and processing

Example of sample sequence in Masslynx

	File Name	File Text	Sample Type	MS File	MS Tune File	Inlet File	Vial	Inj Vol
1	3408L11_01	Bl	Analyte	Multitox	multitox	Multitox	1:1	20
2	3408L11_02	Bl	Analyte	Multitox	multitox	Multitox	1:1	20
3	3408L11_03	STD1	Standard	Multitox	multitox	Multitox	1:2	20
4	3408L11_04	STD2	Standard	Multitox	multitox	Multitox	1:3	20
5	3408L11_05	STD3	Standard	Multitox	multitox	Multitox	1:4	20
6	3408L11_06	STD4	Standard	Multitox	multitox	Multitox	1:5	20
7	3408L11_07	STD5	Standard	Multitox	multitox	Multitox	1:6	20
8	3408L11_08	STD6	Standard	Multitox	multitox	Multitox	1:7	20
9	3408L11_09	STD7	Standard	Multitox	multitox	Multitox	1:8	20
10	3408L11_10	Bl	Analyte	Multitox	multitox	Multitox	1:1	20
11	3408L11_11	Bl	Analyte	Multitox	multitox	Multitox	1:1	20
12	3408L11_12	NK1	Analyte	Multitox	multitox	Multitox	1:9	20
13	3408L11_13	NK2	Analyte	Multitox	multitox	Multitox	1:9	20
14	3408L11_14	Bl	Analyte	Multitox	multitox	Multitox	1:1	20
15	3408L11_15	xxxx a	Analyte	Multitox	multitox	Multitox	1:10	20
16	3408L11_16	xxxx b	Analyte	Multitox	multitox	Multitox	1:11	20
17	3408L11_17	xxxx as	Analyte	Multitox	multitox	Multitox	1:12	20
18	3408L11_18	xxxx bs	Analyte	Multitox	multitox	Multitox	1:13	20
19	3408L11_19	bl	Analyte	Multitox	multitox	Multitox	1:1	20
20	3408L11_20	NK3	Analyte	Multitox	multitox	Multitox	1:9	20
21	3408L11_21	bl	Analyte	Multitox	multitox	Multitox	1:1	20
22	3408L11_22	xxxx a	Analyte	Multitox	multitox	Multitox	1:14	20
23	3408L11_23	xxxx b	Analyte	Multitox	multitox	Multitox	1:15	20
24	3408L11_24	xxxx as	Analyte	Multitox	multitox	Multitox	1:16	20
25	3408L11_25	xxxx bs	Analyte	Multitox	multitox	Multitox	1:17	20
26	3408L11_26	bl	Analyte	Multitox	multitox	Multitox	1:1	20
27	3408L11_27	NK4	Analyte	Multitox	multitox	Multitox	1:9	20
28	3408L11_28	bl	Analyte	Multitox	multitox	Multitox	1:1	20
29	3408L11_29	xxxx a	Analyte	Multitox	multitox	Multitox	1:18	20
30	3408L11_30	xxxx b	Analyte	Multitox	multitox	Multitox	1:19	20
31	3408L11_31	xxxx as	Analyte	Multitox	multitox	Multitox	1:20	20
32	3408L11_32	xxxx bs	Analyte	Multitox	multitox	Multitox	1:21	20
33	3408L11_33	bl	Analyte	Multitox	multitox	Multitox	1:1	20
34	3408L11_34	NK5	Analyte	Multitox	multitox	Multitox	1:9	20
35	3408L11_35	NK6	Analyte	Multitox	multitox	Multitox	1:9	20

When only analyzing for presence of Secalonic Acid in samples it is not necessary to run all standards, since no quantification is done. It is sufficient to analyse Standard 4 together with samples and QC samples.

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Calculations

The MassLynx software QuanLynx is used for calculation of mycotoxin concentration in samle. The Excel spreadsheet is used for calculation of mycotoxin concentration in samle corrected for recovery and dilutions.

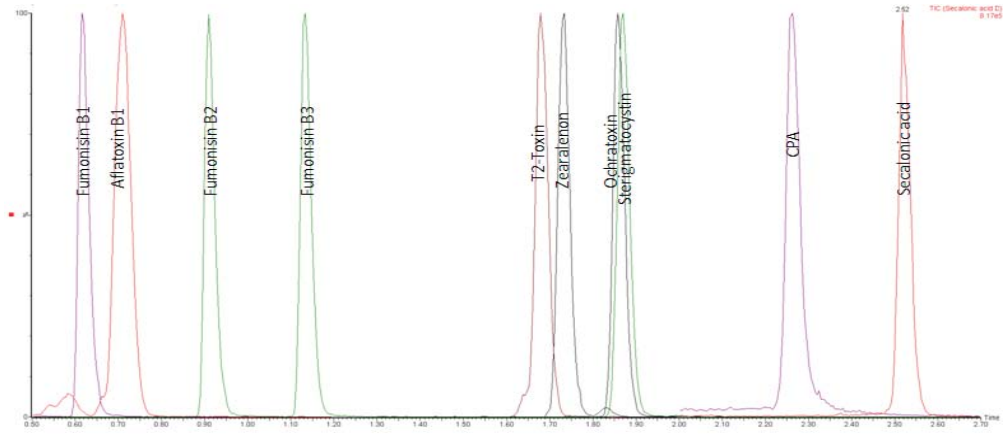
Step	Action
1	Load relevant sample list in MassLynx software
2	Type the exact concentrations of the standards and QC-samples in sample list.
3	Type sample amount (g) or volume (ml) in “User Divisior 1”
4	For unspiked samples: Type “8.40” in “User Factor 1” For spiked samples: Type “8.45” in “User Factor 1” Note: Values in “User Factor 1” corresponds to solvent/sample-ratio obtained during extraction multiplied by final extract dilution.
5	Process relevant Standards, Blanks, QC-samples and samples with relevant process method.
6	Load Excel spreadsheet
7	Select correct sheet according to sample type (i.e “Wheat Bran and Frementation” or “Enzyme conc”.
8	Type following in dedicated excel cells: 1 – Sample registration no 2 – Dilution (if any additional) 3 – Theoretical concentration of spiked sample (in yellow excel cells). Theoretical conc is: Spike conc multiplied by spike volume divided sample amount. Example: $Conc_{CPA} = \frac{0.025ml \times 1500ppb}{1\,\text{ml sample}} = 37.5ppb$ 4 – Concentration of relevant mycotoxins
9	Check if the acceptance criteria for recovery and CV% (spiked and unspiked samples) are met.
10	Type results in LIMS

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Calculations *(continued)*

Chromatogram of standard injection.



Acceptance criteria

Acceptance criteria only apply to the toxins for which analysis is required.

Parameter	Criteria
R^2 for the calibration curves	≥ 0.99
Recovery, Enzyme concentrates	≥ 50%
Recovery, wheat bran & fermentation samples	≥ 25%
CV% (double determinations of spike)	≤ 25%
Concentration of QC sample	± 16% of theoretical concentration

If results are ≥ LOQ for any mycotoxins, then the presence of mycotoxin should be confirmed by the secondary ion transition (i.e the secondary daughter ion). See Component specific settings. Compare ion ratios with ratios of standards.

If in doubt of the identity of a toxin an alternative method could be used for analysis, see Appendix.

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

LOQ

LOQ (ppm):

	Enzyme conc	Fermentation	Wheat bran
CPA	0.01	0.01	0.01
AFB1	0.001	0.001	0.025
OTA	0.001	0.005	0.005
ZEA	0.002	0.002	0.01
T2	0.002	0.002	0.01
STE	0.001	0.001	0.001
SAD	0.01	0.01	0.01
FUM1, 2 & 3	0.001	0.001	0.005

Statement of analytical result

Results for all mycotoxins besides Secalonic acid are reported with two significant differs or <LOQ.

Definition of units: mg/kg (ppm)

Secalonic acid is reported as “Detected”/”Not detected” (> or <LOQ)

Contingencies

All deviations from this SOP should be discussed with the method responsible and should be documented

Archiving

Archive raw data and documents in department archive according to local procedures hyperlinked in.

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Appendix

This appendix contains a list of documents relevant for this method according to local procedures. Also LIMS, Equipment settings, etc. are found in this section.

Equipment	Componentspecific MS settings
LC settings	Validation
LC settings alternative method	LIMS data
MS settings	

Disposal of trash

Nunc tubes should be disposed of in a ziplock bag and put into a white bucket.

Documents relevant for NZDK (PSL):

Document	Link
FARI	EKL-FARI-0064
Validation	2011-27024-01
Data capture template	EKL-PR-035
Preparation of Eluent A	PSL-MR-0199
Preparation of Eluent C	PSL-MR-0200
Preparation of stock solutions for standards and spike	EKL-SF-0062
Preparation of stock solutions for QC sample	EKL-SF-0103
Preparation of working standard solution	EKL-RA-0013
Preparation of QC sample	EKL-RA-0042
Calculations	EKL-RA-0043
UPLC system	EKL-TE-4063.01-D
Quattro Premier XE MS and Xevo-TQ-MS	EKL-TE-4060.01-D
Masslynx	EKL-SP-3110.01-D

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Equipment

LC system	Ultra performance liquid chromatography (UPLC) (Appendix)
MS/MS	Primary (DK): Quattro Premier XE (Appendix) Backup (DK): Xevo-TQ-MS (Appendix)
Software	MassLynx
Vials	Glass vials

LC-settings

Run time	6.99 min
Flow	0.7 mL/min
Column temperature	60°C
Autosampler temperature	Max. 5-15°C
Injection volume	20 µL
Gradient type	6
Needle type	Peak
Loop	20 µl
Injection mode	Partial loop with overfill or full loop.
Column storage	Water/ACN (50:50)
Column type	Material: BEH-C18, length: 50mm, size:1.7µm, ID: 2.1 mm

Gradient table:

Time [min]	A1-eluent (A) [%]	B1-eluent (C) [%]
0.00	80	20
3.00	50	60
3.50	10	90
4.00	80	20
6.05	80	20
7.00	80	20
7.1	50	50

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

LC-settings alternative method

If additional confirmation of mycotoxins is needed, use alternative LC-settings and parameters. The methods are all called mycotoz_alt in Masslynx

Run time	11.5 min
Flow	0.35 mL/min
Column temperature	60°C
Autosampler temperature	Max. 5-15°C
Injection volume	20 µL
Gradient type	6
Needle type	Peak
Loop	20 µl
Injection mode	Partial loop with overfill or full loop.
Column storage	Water/ACN (50:50)
Columns type	Material: CSH-C18, length: 100mm, size:1.7µm, ID: 2.1

Gradient table:

Time [min]	A1-eluent (A) [%]	B1-eluent (C) [%]
0.00	99	1
7.5	40	60
8	5	95
9	5	95
9.5	99	1
11.99	99	1
12.50	50	50

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

MS settings

Detector model	MS/MS (Quattro Premier XE)	MS/MS (Xevo)
General settings: **Source**	• Capillary: 1.00 kV • Cone: Component specific • Extractor: 2 V • RF-lens: 0.5 V • Source temp: 125° C • Desolvation temp: 350° C • Cone gas flow: 30 L/hour • Desolvation gas flow: • 850 L/hour	• Capillary: 1.00 kV • Cone: Component specific • Extractor: 2.5 V • RF-lens: • Source temp: 150° C • Desolvation temp: 650° C • Cone gas flow: 100 L/hour • Desolvation gas flow: • 1100 L/hout

Detector model	MS/MS (Quattro Premier XE)	MS/MS (Xevo)
General settings: **MS-file**	• Inter-channel delay: 0.02 sek. • Inter-scan delay: 0.02 sek. • Repeats: 1 • Span: 0.2 • Dwell time: Component specific • For other parameters see the detector component-specific settings	• Inter-channel delay: Auto. • Inter-scan delay: Auto. • Repeats: • Span: 0.2 • Dwell time: Auto. •
Described in SOP	EKL-TE-4060.01-D and EKL-SP-3110.01-D.	EKL-TE-4060.01-D and EKL-SP-3110.01-D.

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Component-specific settings

Component	Settings				
		Parent > Daughter:	Cone:	Coll Energy:	Dwell (secs)
Aflatoxin B_1 $[M+H]^+$ (Rt = 0.60-0.70 min)	**Premier XE**	313.2 > 241.2 313.2 > 285.1	40 V	38 eV 22 eV	0.1 0.1
	Xevo	313.1 > 213.2 313.1 > 241.1 313.2 > 285.3	47 V	42 eV 36 eV 21 eV	0.03 0.03 0.03
Zearalenon $[M-H]^-$ (Rt = 1.60-1.70 min)	**Premier XE**	317.2 > 130.9 317.2>174.98	40 V	32 eV 24 eV	0.04 0.062
	Xevo	317.2 > 130.9 317.2 >174.98	40 V	32 eV 24 eV	0.018 0.018
T-2 toxin $[M+Na]^+$ (Rt = 1.60-1.70 min)	**Premier XE**	489.3 > 245.2 489.3 > 327.2	38 V	24 eV 24 eV	0.036 0.062
	Xevo	489.3 > 245.1 489.3 > 327.2	38 V	26 eV 24 eV	0.018 0.018
Ochratoxin A $[M+H]^+$ (Rt = 1.75-1.85 min)	**Premier XE**	404.3 > 239.1 404.3 > 358.2	22 V	23 eV 13 eV	0.048 0.048
	Xevo	404.3 > 220.97 404.3 > 239.1	20 V	38 eV 28 eV	0.018 0.018

Continued on next page

Mycotoxins by LC-MS/MS, *Continued*

Component-specific settings

Component	Settings

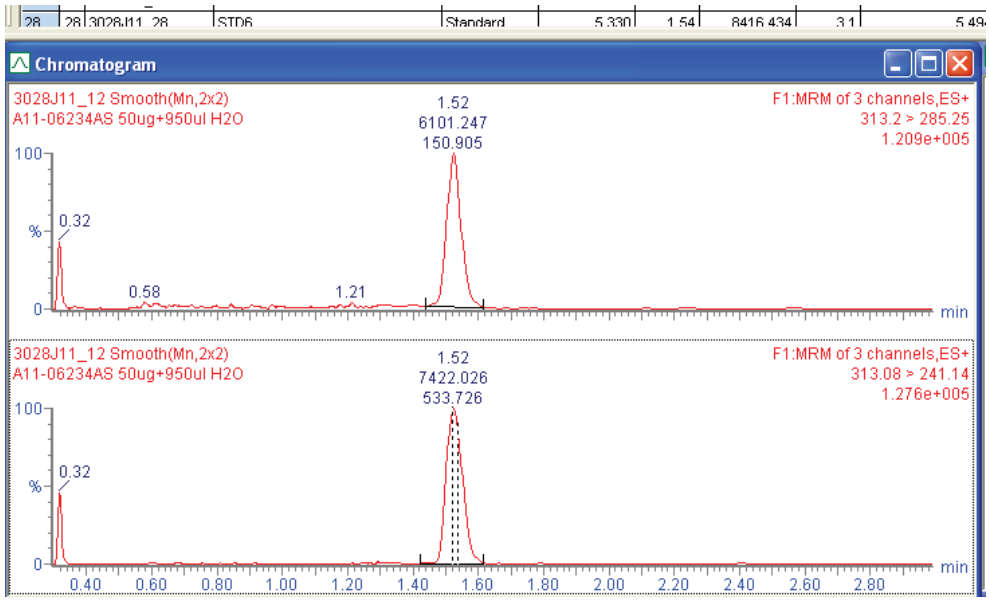
		Parent > Daughter:	Cone:	Coll Energy:	Dwell (secs)
Sterigmatocystin $[M+H]^+$ (Rt = 1.75-1-85 min)	**Premier XE**	325.2 > 281.1	40 V	35 eV	0.031
		325.2 > 310.1		25 eV	0.031
	Xevo	325.3 > 281.1	38 V	36 eV	0.018
		326.3 > 310.1		24 eV	0.018
CPA $[M+H]^+$ (Rt = 2.15-2.25 min)	**Premier XE**	337.3 > 182.2	28 V	22 eV	0.2
		337.3 > 196.2		22 eV	0.2
	Xevo	337.3 > 182.2	24 V	44 eV	0.038
		337.3 > 196.2		26 eV	0.038
Secalonic Acid D $[M+H]^+$ (Rt = 2.40-2.50 min)	**Premier XE**	639.3 > 183.2	35 V	35 eV	0.2
		639.3 > 561.2		35 eV	0.2
	Xevo	639.3 > 150.97	40 V	50 eV	0.038
		639.3 > 561.2		24 eV	0.038
Fumonisin B1 $[M+H]^+$ (Rt = 0.68-0.88 min)	**Premier XE**	722.4 > 334.2	40 V	40 eV	0.1
	Xevo	722.5 > 352.4	48 V	34 eV	0.032
Fumonisin B2+3 $[M+H]^+$ (Rt = 1.11-1.25 min)	**Premier XE**	706.4 > 318.3	55 V	40 eV	0.1
	Xevo	706.5 > 354.4	46 V	32 eV	0.018

Continued on next page

Mycotoxins by LC-MS/MS, Continued *Continued*

Component-specific settings *(continued)*

Example of chromatogram of Aflatowin showing both primary and secondary ion transition



Validation data

Parameter	Result
Repeatability $_{\text{LOQ level}}$	≤ 30%
Intermediate precision $_{\text{LOQ level}}$	≤ 35%
Linearity (r^2)	≥ 0.995
Accuracy $_{\text{Enzyme concentrates}}$	≥ 80%
Accuracy $_{\text{Fermentation}}$	≥ 50%
Accuracy $_{\text{Wheat bran}}$	≥ 35%

Continued on next page

Mycotoxins by LC-MS/MS, Continued*Continued*

Function	Data
Group name:	TX
Method name	TX-AFLA B1 TX-AFLAB1(W_V) TX-CPA TX-CPA(W_V) TX-OCHRA A TX-OCHRA A(W_V) TX-STERIGMA TX-STERIGMA(W_V) TX-T2 TOX TX-T2 TOX(W_V TX-ZEARAL TX-ZEARAL(W_V) TX-FUM B1; TX-FUM B"; TX-FUM B3 TX-SECALON A
Unit	mg/kg and mg/L
Decimals (AFLA, T2, ZEA, FUM, OTA, STE)	4
Decimals (CPA, SECA)	3
Min - Max estimate (AFLA, FUM, OTA, STE)	0.001-0.1
Min - Max estimate (ZEA, T2)	0.002-0.1
Min - Max estimate (CPA, SECA)	0.01-0.1
Quality class and CVsingle det.	10-24%
Method Control req.	N
Automatic approval	Y
QC sample	N

Revision

Version 2.0: Changes in order of toxins, so that they are the same throughout the document. Whirlmixing of samples only necessary for solid samples. When only analyzing for Secalinc Acid D it is added that sample set could be run with only one standard and not a whole standard curve. Recommended dilution volumes for different sample types added. 2012-01-30 JRzK.

LC-settings for alternative method changed. 2012-03-16 RFHa

Detection of production strains

Scope

All Novozymes QC laboratories involved in analysis of samples from Novozymes production and GLP studies.

Content

Principle

The analysis is performed by spread plating or enrichment of a known amount of the sample. By spread plating, the colonies on the test-plates are compared morphologically with the colonies of the reference strain and suspect colonies be verified as described in BD 002-IN or on a corresponding Analytical Directions (where applicable).

The production strain

Is defined as the microorganism used for fermentation of a given Novozymes product. Agar media and incubation conditions used for detection of a specific production strain are listed in BD 002-IN.

Continued on next page

Detection of production strains, *Continued*

Principle *(continued)*

The reference strain

is defined as an isolate of the production strain used in the laboratory as a reference during the analysis.

Strains not listed in BD 002-IN are detected according to specific Analytical Directions.

Analytical Direction (AD)

A description of the reference strain, agar type, incubation period and temperature to be used for the analysis of production strains that are not yet included in the BD 002-IN. It is prepared and approved by the Method Responsible Scientist. Analytical Directions are typically used in connection with GLP studies.

- When analyzing samples from Novozymes production, the detection is carried out by spread plating of 0.1 g or 0.1 ml of sample.
- When analyzing samples from GLP studies, the detection is carried out by spread plating or enrichment of 1 g of sample acc. to the specific AD.

Detection of morphologically typical colonies (compared with the reference strain) indicates the presence of the production strain.

Continued on next page

Detection of production strains, *Continued*

Definition of units

When analyzing samples from **Novozymes production**, the result is stated as:

- **DET** (The production strain detected in 0.1 g or ml)
- **ND** (The production strain not detected in 0.1 g or ml)

When analyzing samples from **GLP studies**, the result is stated as:

- **DET** (The production strain detected in 1 g) *or*
- **ND** (The production strain not detected in 1 g)

IMPORTANT: When detected, the approximate number of production strain (CFU) per g or ml is stated provided the analysis was performed using spread plating.

Sample type

All Novozymes samples from production and GLP studies.

Detection limit

The detection limit of this method is dependent on the sample volume and the dilution in use.

Sample volume	Size and number of agar plates	Dilution	Detection limit
1 ml	14 cm (4 plates)	10^{-1}	10 CFU / (g or ml)
10 ml	14 cm (20 plates)	10^{-1}	1 CFU / (g or ml)

Equipment and materials

- Balance (± 0.1 g)
- Refrigerator
- Magnetic stirrer
- Petri dishes (14 cm and 9 cm)
- Suitable sterile pipettes for transfer of 10 ml, 1 ml (4x0.25 ml) and 0.25 ml
- Sterile Drigalski spatula
- Incubator (relevant incubation temperatures are listed in BD 002-IN)

Continued on next page

Detection of production strains, *Continued*

Media and reagents for Bacterial strains

Dilution buffer: Tween buffer 4%, 90 ml (if necessary, with a magnet) prepared acc. to EB-ME-0052 (Danish sites: MSA-SUB-FS-0530).

Agar media:

Abbreviation	Full name	Prepared acc. to "Danish sites"	Prepared acc. to EB Media direction
AT-2	AT-2 agar	MSA-SUB-FS-0615	EB-ME-0001
B-TSA	Basic Tryptic Soy Agar	MSA-SUB-FS-0572	EB-ME-0055
B-TSA w. CAM	Basic Tryptic Soy Agar w/wo Chloramphenicol (CAM). *NOTE*: The addition of CAM is optional	MSA-SUB-FS-0572	EB-ME-0056
Schaeffers	Schaeffers agar	MSA-SUB-FS-0488	EB-ME-0036
Sch.starch	Schaeffers agar with 1% starch	MSA-SUB-FS-0232	EB-ME-0037
TSA with Skim milk	Tryptic Soy Agar with 1 % skim milk		EB-ME-0038
TSApH9 with Skim milk	Tryptic Soy Agar at pH 9 with 1% skim milk	MSA-SUB-FS-0660	EB-ME-0070
TBX w. AMP	Chromocult®TBX agar with ampicillin (100 mg/l)	MSA-SUB-FS-0452	EB-ME-0066
TSA	Tryptic Soy Agar	MSA-SUB-FS-0260	EB-ME-0041
TSA w. KANA	Tryptic Soy Agar w/wo kanamycin (Kana) *NOTE*: The addition of KANA is optional	MSA-SUB-FS-0243	EB-ME-0058

Continued on next page

Detection of production strains, *Continued*

Media and reagents for Fungal strains

Dilution buffer: Tween buffer 4%, 90 ml (if necessary, with a magnet) prepared acc. to EB-ME-0052 (Danish sites: MSA-SUB-FS-0530).

Agar media:

Abbreviation	Full name	Prepared acc. to "Danish sites"	Prepared acc. to EB Media direction
Cove-T-2	Cove-T-2 agar	MSA-SUB-FS-0320	EB-ME-0013
DG-18	DG-18 agar	MSA-SUB-FS-0132	EB-ME-0017
Phytate	Phytate agar		EB-ME-0028
Sch.starch	Schaeffers agar with 1% starch	MSA-SUB-FS-0232	EB-ME-0037
YPG w/wo Tetracycline	YPG agar w/wo tetracycline. *NOTE:* The addition of tetracycline is optional	MSA-SUB-FS-0274	EB-ME-0044
YPSS w/wo Tetracycline	YPSS agar w/wo tetracycline. *NOTE*: The addition of tetracycline is optional	MSA-SUB-FS-0278	EB-ME-0045
YSG	Yeast/Soy Peptone/Glucose	MSA-SUB-FS-0664	EB-ME-0071
PDA	Potato Dextrose Agar	MSA-SUB-FS-0380	EB-ME-0075
MEA **(NZIN only)**	Malt Extract Agar Base w/mycological peptone		EB-ME-0077

Continued on next page

Detection of production strains, *Continued*

Safety

It is the responsibility of the laboratory leader, that all personnel are aware of the correct handling of enzymes, reagents and microorganisms.

CAUTION: Cryo-tubes with fungi production strains may contain 5% DMSO which is a health hazardous substance.

NOTE: Agar plates containing fugal production strain may **only** be opened in safety cabinet.

Sample preparation

The samples are prepared as follows:

Step	Action
1	Transfer 10 g of solid sample or 10 ml of liquid sample to 90 ml Tween buffer 4%.
2	Immediately homogenize the sample by stirring or by shaking. Solid samples are homogenized on a magnetic stirrer for app. 20 minutes.

IMPORTANT: All enzyme products must be analyzed from a 10^{-1} dilution due to possible inhibition of microorganisms in undiluted enzyme.

TIP: Non-enzyme liquid samples (*e.g*. CIP-samples) are analyzed undiluted.

TIP: Further 10-fold dilutions of any sample type can be prepared with Tween buffer 4% as needed.

Continued on next page

Detection of production strains, *Continued*

Plating

The relevant agar plates and incubation conditions (time and temperature) are listed in BD 002-IN or the relevant AD.

Step	Action
1	Prepare the test plates: • When analyzing samples from **Novozymes production**: Transfer 1 ml from the 10^{-1} dilution onto the surface of 4 relevant agar plates (14 cm) with app. 0.25 ml on each plate. • When analyzing samples from **Tox batches (GLP)**: Analyze according to the relevant AD.
2	Prepare the control plates: **Positive Product control (PPC) plates** (spread plating): • Transfer 0.25 ml from the 10^{-1} dilution onto the surface of one relevant agar plate (14 cm), and streak the bacterial reference strain or point inoculate the fungal production strain onto the inoculated plate. **Positive control (PC)** (spread plating): • Streak the bacterial reference strain or point inoculate the fungal strain onto another agar plate (not inoculated with sample). **Positive control** (enrichment + spread plating): • Prepare a positive control as described in the relevant AD
3	Leave the plates on the table until the sample has been soaked into the agar.

NOTE: Incubation time is just a guideline. The plates can be read whenever the reference strain demonstrates typical colonies on the positive control plates.

REFERENCE: EB-SP-5006.02

Continued on next page

Detection of production strains, *Continued*

Reading

The colonies on the test-plates are compared morphologically with the colonies of the reference strain.

If …	Then …
No suspect colonies are observed on the test-plates …	The result is stated as: **ND** (the production strain is Not Detected)
Too heavy growth of accompanying flora which makes it impossible to decide whether the production strain is present or not……	The result is stated as: **NOT READABLE**
Suspect colonies are observed on the test-plates …	The suspect colonies are verified. See Verification. *NOTE*: If possible, the number of colonies is assessed.

IMPORTANT: The reference strain must grow on both positive control plates. If not, the test must be repeated.

Verification

The reference strain and minimum one colony of each type of suspect colonies from the test plates are streaked or point inoculated onto one or more of the agar plates (9 cm or 14 cm) listed in BD 002-IN (column "Verification") or in the AD.

Inoculation and reading of these agar media are described below. The plates are incubated as listed in BD 002-IN (column "Verification") or in the AD. If necessary, these media can be supplemented with other agar media, *e.g.* the agar medium used for the detection.

NOTE: Incubation time is just a guideline. The plates can be read whenever the reference strain demonstrates typical colonies on the positive control plates.

NOTE: If verification is performed on Schaeffers agar with starch, the Lugol's iodine solution can be used for verification of amylase activity - Merck Cat. No. 109261 or equivalent reagent.

Continued on next page

Detection of production strains, *Continued*

AT-2 agar

Detection of **pullulanase** activity:

	Description
Principle	Pullulanase-producing strains degrade the amylopectin in the agar. Thus, staining the plate with Lugol's solutions results in blue zones (haloes) surrounding the colony of the isolate.
Inoculation	Point inoculation
Reading	Colonies of the isolate are compared morphologically with the colonies of the reference strain. The surface of the plates is carefully flooded with Lugol's solution. Blue zones surrounding the colony in a reddish-brown medium indicate pullulanase activity. *NOTE:* If the production strain produces amylase in addition to pullulanase, clearing zone will surrounds the colony. Between the clearing zone and the reddish-brown medium a narrow blue zone might be seen.

Cove-T-2 agar

Detection of ***amdS*-transformed** fungi:

	Description
Principle	GMO strains transformed with the marker *amdS* grow well on the agar, while other strains grow poorly or not at all.
Inoculation	Point inoculation.
Reading	Colonies of the isolate are compared morphologically with the colonies of the reference strain. Vigorous growth on Cove-T-2 indicates presence of an *amdS*-transformed strain.

Continued on next page

Detection of production strains, *Continued*

DG-18 agar Comparison of **morphology** of fungi:

	Description
Principle	DG-18 is a general growth medium for Fungi. The agar is used for comparison of morphology of fungal isolates with the reference strain.
Inoculation	Point inoculation
Reading	Colonies of the isolate are compared morphologically with the colonies of the reference strain.

MEA agar Comparison of **morphology** of fungi:

	Description
Principle	MEA is a media recommended for detection, isolation and enumeration of yeast and molds. The agar is used for comparison of morphology of fungal isolates with the reference strain.
Inoculation	Point inoculation
Reading	Colonies of the isolate are compared morphologically with the colonies of the reference strain.

PDA agar Comparison of **morphology** of fungi:

	Description
Principle	PDA is recommended for the isolation and enumeration of yeast and molds. The agar is used for comparison of morphology of fungal isolates with the reference strain.
Inoculation	Point inoculation
Reading	Colonies of the isolate are compared morphologically with the colonies of the reference strain.

Continued on next page

Detection of production strains, *Continued*

Phytate agar

Detection of **phytase** activity:

	Description
Principle	Phytase-producing strains degrade phytate in the agar. Thus, clearings zone (halo) that surround the colony of the isolate indicates phytase activity.
Inoculation	Point inoculation
Reading	Colonies of the isolate are compared morphologically with the colonies of the reference strain. Before inoculation the plates are opaque but any presence of phytase activity will result in clearings zone (halo) that will surround the colony.

Schaeffers agar

Sporulation test (*Bacillus* spp.):

	Description
Principle	Schaeffers agar induces sporulation of wild type strains due to nutrient limitation of the media. But the production strains show no sporulation on Schaeffers agar after incubation for 2-3 days, since they are sporulation deficient.
Inoculation	Streaking
Reading	Colonies of the isolate are compared morphologically with the colonies of the reference strain. The colonies are examined by phase-contrast microscopy for presence of spores which can be observed as luminous cells. The production strain shows no or very little sporulation after incubation for 2-3 days.

Continued on next page

Detection of production strains, *Continued*

Schaeffers starch agar

Detection of **amylase** activity (all isolations) / **sporulation** test (*Bacillus* spp.):

	Description
Principle	***Bacillus spp.***: Schaeffers agar induces sporulation of wild type *Bacillus* strains, whereas the *Bacillus* production strains, which are sporulation defect, show no or very little sporulation on Schaeffers agar after incubation for 2-3 days. ***Bacillus spp.* & Fungi**: Amylase producing strains degrade the starch in the agar. Thus, clearings zone (halo) that surround the colony of the isolate indicates amylase activity.
Inoculation	Point inoculation
Reading	Colonies of the isolate are compared morphologically with the colonies of the reference strain. ***Bacillus spp.:*** The colonies are examined by phase-contrast microscopy for presence of spores which can be observed as luminous cells. The production strain shows no or very little sporulation after incubation for 2-3 days. ***Bacillus spp.* & fungi:** The surface of the plates is carefully flooded with Lugol's solution. A clearings zone that surround the colony in a blue (dark blue) medium indicates amylase activity.

Skim milk agar

Detection of **proteolytic** activity:

	Description
Principle	Protease-producing strains degrade the skim milk in the agar. Thus, clearings zone (halo) that surround the colony of the isolate indicates presence of protease.
Inoculation	Point inoculation
Reading	Colonies of the isolate are compared morphologically with the colonies of the reference strain. The plates are generally opaque, so the presence of clearing zone (halo) that surround the colony of the isolate will, therefore, indicates proteolytic activity.

Continued on next page

Detection of production strains, *Continued*

Interpretation of results

If …	Then …
suspect colonies are verified as the production strain…	the result is stated as: **DET** (Production strain Detected). *NOTE:* The number of colonies on the original plate is estimated (if possible) and the CFU/g or ml is stated in the LIMS "Notes" field. *IMPORTANT:* Contact the local method responsible scientist and inform both the QCC-cor. and the Dept. Manager of the submitter by mail.
there are no suspect colonies *or* suspect colonies cannot be verified as the production strain…	the result is stated as: **ND** (Production strain is Not Detected)
it is not possible to read the test plates due to heavy growth of accompanying flora…	the analysis is "Complete without result" in LIMS and stated as: **NOT READABLE**

Accuracy and precision

It has not been determined, but it will depend on the type of enzyme product as well as the detection/verification method.

Filing

All documentation should be filed in accordance with the local filing SOP.

Contingencies

All systematic deviations from this SOP should be discussed with the Method Responsible Scientist and should be documented.

References

BD 002-IN (**must** apply for access – limited access).

Revision

Generally, the reference BD 001-IN-000 has been changed to BD 002-IN throughout the document.
pp. 1: Added section "Content".
pp. 3: In section "Definition of units", it is emphasized that estimation of the approximate number of production strain / g or ml is only done when the analysis was performed using spread plating.

Continued on next page

Detection of production strains, *Continued*

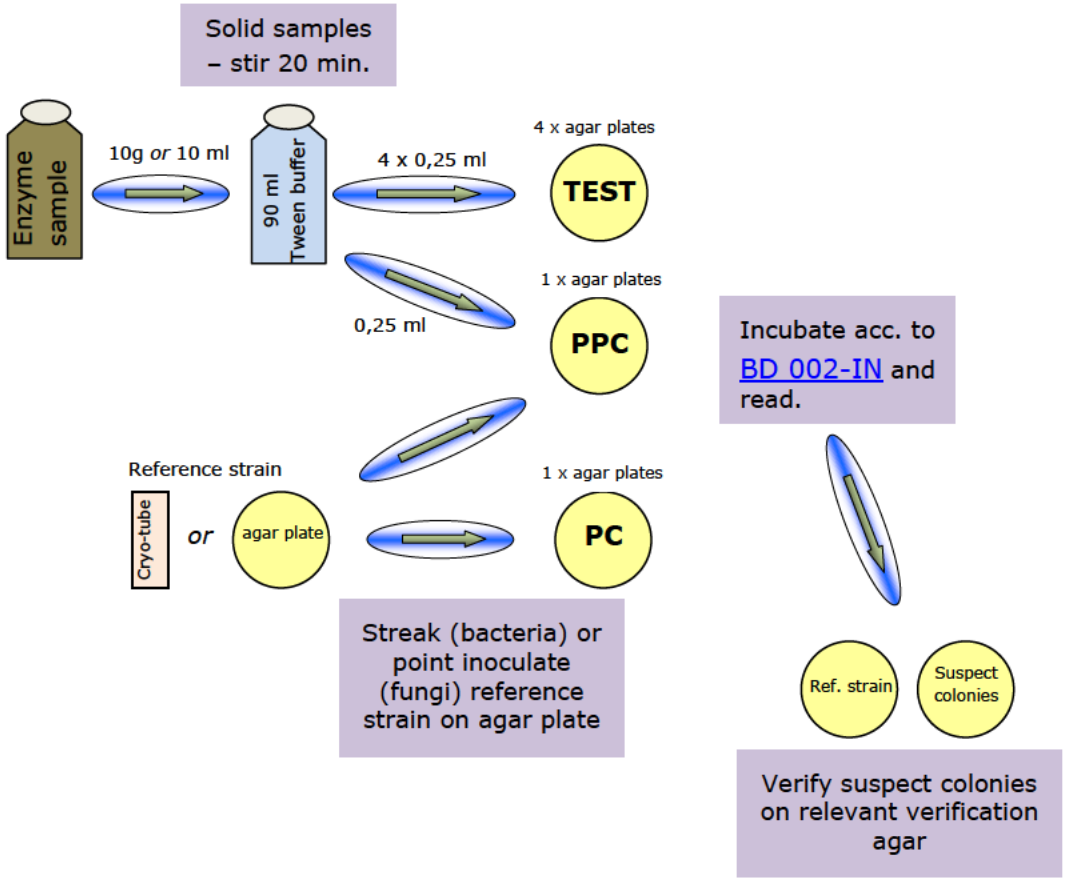
Revision *(continued)*

pp. 3: In section "Sample type" the sentence was changed.
pp. 4+5: Added Danish media directions when available. Furthermore, following agar media were added to the tables: TSApH9 Skim Milk, MEA and PDA.
pp. 6: In section "Safety" added that 5% DMSO is a health hazardous substance and that agar plates containing fungal production strain may only be opened in safety cabinet.
pp. 6: In section "Sample preparation" added that non-enzyme liquid samples are analyzed undiluted and that further10-fold dilutions of enzyme samples can be prepared with Tween buffer 4%.
pp. 7: In section "Plating" added preparation of positive control when detection is carried out using an enrichment step and a note that incubation time is just a guideline. Furthermore, added a reference.
pp. 8: In section "Verification" specified that minimum one colony of each type of suspect colonies must be verified and added that vendor Lugol's iodine solution may be used instead of a solution prepared according to EB-ME-0021.
pp. 10: Added following sections "MEA agar" and "PDA agar" to describe the principle of the verification.
pp. 13: The title of section "Calculation" has been changed to "Interpretation of results" and has been extensively revised.
pp. 14: Section "Accuracy and Precision" has been extensively revised to reflect the real content of this section instead of describing the detection limit (already described on pp. 3).
pp. 14: The title of section "Archiving" has been changed to "Filing".
pp. 15: Added Flow chart.
Other minor editorial changes.
This document replaces PSL-SM-1200.01-D version 11.
Revised by CSLC.

Continued on next page

Detection of production strains, *Continued*

Flow Chart Click to read section.



Appendix 4

Documentation regarding the manufacturing process

1. Statement on compliance of Good Manufacturing Practices, Food
2. ISO 9001:2015 certificate

11 June 2018
Statement no. 401.15

To Whom It May Concern

Statement on Good Manufacturing Practice - GMP
- general description of production, control and hygiene

Novozymes A/S is a manufacturer of enzymes used in the food industry. We hereby certify that: The products are produced according to good manufacturing practices for manufacturing, packing, or holding human food in order to prevent serious food hazards. Furthermore, our documented quality system is ISO 9001[1] certified by Bureau Veritas Certification, accredited by UKAS. The quality system includes:

- Production operations are conducted in accordance with adequate sanitation principles.
- HACCP plan. Critical control points (CCPs) are identified and controlled, and the products are released if in compliance with these requirements.
- Critical measuring equipment is identified and calibrated at regular intervals.
- Instructions on cleaning of equipment, utensils and rooms are established and cleaning is documented.
- The personnel is trained in hygienic practices in order to prevent contamination of products and equipment.
- The personnel is trained in the quality system.
- The buildings and equipment are monitored periodically with special reference to maintenance.
- The production of our food enzymes complies with EC regulation 852/2004/EC, including amendments, on *the hygiene of foodstuffs.*
- The packaging materials used for our food enzyme products comply with EC regulation 1935/2004/EC, and related legislation including amendments on materials and articles intended to come into contact with foodstuffs.
- The production is under control of and inspected by the authorities according to EC regulation 882/2004/EC, including amendments, on *the official control of foodstuffs* as interpreted and implemented in Danish legislation.

[1]The scope of the 9001 certificate is: Development, Production and Sales of Biopolymers and Industrial Enzymes.

Novozymes A/S
Krogshoejvej 36
2880 Bagsvaerd
Denmark

Tel. +45 4446 0000
Fax +45 4446 9999

BUREAU VERITAS
Certification

Novozymes A/S

Krogshøjvej 36, 2880 Bagsværd, Denmark

This is a multi-site certificate. Additional site details are listed in the appendix to this certificate.

Bureau Veritas Certification Holding SAS – UK Branch certifies that the Management System of the above organization has been audited and found to be in accordance with the requirements of the management system standards detailed below.

Standard

ISO 9001:2015

Scope of certification

Development, Production and Sales of Industrial Enzymes.

Original cycle start date:	**25 March 1996**
Expiry date of previous cycle:	**NA**
Certification/Recertification Audit date:	**NA**
Certification/Recertification cycle start date:	**26 March 2018**

Subject to the continued satisfactory operation of the organization's Management System, this certificate expires on: **25 March 2021**

Certificate No.: DK008854 Version: 1 Revision date: **08 March 2018**

MIX Paper from responsible sources FSC® C074440

Certification body address: *5th Floor, 66 Prescot Street, London, E1 8HG, United Kingdom*
Local Office: *Oldenborggade 25-31, 7000 Fredericia, Denmark*

UKAS MANAGEMENT SYSTEMS 0008

Further clarifications regarding the scope of this certificate and the applicability of the Management System requirements may be obtained by consulting the organization.
To check this certificate validity, please call **(+45) 77 311 000.**

BUREAU VERITAS
Certification

Novozymes A/S

Standard

ISO 9001:2015

Scope of certification

Development, Production and Sales of Industrial Enzymes.

Site Name/location:	Site Addition Date:	Site Address:	Site Scope:
Novozymes A/S (Head Office)	25-03-1996	Krogshøjvej 36, 2880 Bagsværd, Denmark	Development, Production and Sales of Industrial Enzymes.
Novozymes A/S	25-03-1996	Hillerødgade 31 & 42, 2200 København N, Denmark	Development, Production and Sales of Industrial Enzymes.
Novozymes A/S	25-03-1996	Hallas Allé 1, 4400 Kalundborg, Denmark	Development, Production and Sales of Industrial Enzymes.
Novozymes /China) Biotechnology Co., Ltd.	25-03-1996	No. 150 Nanhai Road, TEDA, Tianjin, P.R.China	Design and Development, Production and Service of Industrial Enzymes.
Suzhou Hongda Enzyme Co., Ltd.	25-03-1996	Shaxi Town, Taicang City, Suzhou City, Jiangsu Province, P.R.China	Production and Service of Industrial Enzymes.

Certificate No.: DK008854 Version: 1 Revision date: **08 March 2018**

Certification body address: *5th Floor, 66 Prescot Street, London, E1 8HG, United Kingdom*
Local Office: *Oldenborggade 25-31, 7000 Fredericia, Denmark*

Further clarifications regarding the scope of this certificate and the applicability of the Management System requirements may be obtained by consulting the organization.
To check this certificate validity, please call **(+45) 77 311 000.**

UKAS MANAGEMENT SYSTEMS
0008

MIX Paper from responsible sources FSC® C074440

BUREAU VERITAS
Certification

Novozymes A/S

Standard

ISO 9001:2015

Scope of certification

Site Name/location:	Site Addition Date:	Site Address:	Site Scope:
Novozymes (China) Investment Co., Ltd.	25-03-1996	No. 14 XinXi Road, Shangdi Zone, Haidian District, Beijing, P.R.China	Design and Development and Sales of Industrial Enzyme. Biological Products for Wastewater Treatment, Household and Industrial Cleaning and Aquaculture Treatment.
Novozymes USA	25-03-1996	77 Perry Chapel Church Road, Franklinton, North Carolina, 27525-0576, USA	Development, Production and Sales of Industrial Enzymes.
Novozymes Nebraska	25-03-1996	600 S. 1st Street, NE 68008, Blair, USA	Development, Production and Sales of Industrial Enzymes.
Novozymes Brazil	25-03-1996	Rua Professor Francisco Ribeiro 683, CEP 83707-660, Bairro Barigüi, Araucária – Paraná, Brazil	Development, Production and Sales of Industrial Enzymes.

Certificate No.: DK008854 Version: 1 Revision date: **08 March 2018**

Certification body address: *5th Floor, 66 Prescot Street, London, E1 8HG, United Kingdom*
Local Office: *Oldenborggade 25-31, 7000 Fredericia, Denmark*

Further clarifications regarding the scope of this certificate and the applicability of the Management System requirements may be obtained by consulting the organization.
To check this certificate validity, please call **(+45) 77 311 000.**

UKAS MANAGEMENT SYSTEMS
0008

MIX Paper from responsible sources FSC® C074440

BUREAU VERITAS
Certification

Novozymes A/S

Standard

ISO 9001:2015

Scope of certification

Site Name/location:	Site Addition Date:	Site Address:	Site Scope:
Novozymes South Asia Private Limited	25-03-1996	Genisys Building, Plot No. 32, 47-50, EPIP Area Bangalore, 560066 Karnataka, India	Development and Sales of Industrial Enzymes.
Novozymes South Asia Pvt Limited	25-03-1996	Survey No: 193, Hoody Village. Whitefield Road, 560048 Bangalore, India	Production and Sales of Industrial Enzymes.
Novozymes South Asia Pvt. Ltd.	08-03-2018	Plot No.A-1, Patalganga-Borivali Industrial Area, Patalganga, Khalapur, Raigad – 410202, Maharashtra, India	Development, Production and Sales of Industrial Enzymes.

Certificate No.: DK008854 Version: 1 Revision date: **08 March 2018**

Certification body address: *5th Floor, 66 Prescot Street, London, E1 8HG, United Kingdom*
Local Office: *Oldenborggade 25-31, 7000 Fredericia, Denmark*

Further clarifications regarding the scope of this certificate and the applicability of the Management System requirements may be obtained by consulting the organization.
To check this certificate validity, please call **(+45) 77 311 000.**

UKAS MANAGEMENT SYSTEMS
0008

Appendix 5

Safety documentation

1. Glucoamylase produced by 41SaM2-54. Assessment of sequence homology to known toxins and allergens. Novozymes Report No.: 2018-06838

2. Allergen/Toxin Risk Assessment. Novozymes Report No.: 2018-07032

3. Summary of toxicity data. Amyloglucosidase. Novozymes Report No.: 2013-10425

4. Amyloglucosidase, PPY 32789: Test for mutagenic activity with strains of *Salmonella typhimurium* and *Escherichia coli*. Novozymes Study No. 20118069.
   Novozymes Report no.: 2011-30186

5. Amyloglucosidase, PPY32789: Induction of micronuclei in cultured human peripheral blood lymphocytes. Covance Laboratories Study No. 8259272.
   Novozymes Report No.: 2012-10374

6. Amyloglucosidase, PPY24900: Toxicity Study by Oral Administration to CD Rats for 13 Weeks. Huntingdon Life Sciences Study No. NVZ0028/053878.
   Novozymes Report No.: 2006-46592

Glucoamylase produced by 41SaM2-54
Assessment of sequence homology to known toxins and allergens

Esben Friis
Luna# 2018-06838-01

May 17, 2018

Contents

---

1 Sequence homology of glucoamylase from 41SaM2-54 to known toxins

Uniprot database

Protein sequences that contain the word toxin in the description field were extracted from UNIPROT (Database date: 2018-04-25). This database contains entries from SWISSPROT and TREMBL. 108831 entries were found. Each of the sequences was placed in its uniquely named Fasta file. The glucoamylase from 41SaM2-54 sequence was placed in a separate file: glucoamylase.fasta. The script in appendix A was used to invoke the sequemce alignment program ClustalW 2.0.10 to align each sequence to glucoamylase from 41SaM2-54. A summary file containing the length of each sequence and number of identical residues is also created. From this, the identity percentage to the glucoamylase from 41SaM2-54 sequence or the compared toxin sequence is calculated, whichever is longest. This is chosen because the toxin sequences have many different lengths, both much shorter and much longer than the glucoamylase from 41SaM2-54 sequence. By always using the longest sequence, artificial high scores from very short or very long toxins are avoided. The largest homology encountered was 16.2%, indicating that the homology to any toxin sequence in this databas is indeed random and very low. Results with more than 10% identity, or the 1000 results with largest identity, are shown in appendix B.

2 Allergen analysis of glucoamylase from 41SaM2-54

2.1 Allergen Databases

- `http://allergenonline.org`. This is the home page of the The Food Allergy Research and Resource Program (FARRP) allergen protein database. The present report use data downloaded 2018-01-18. Appendix D shows a list.

2.2 Analyses

1. **35% identity over 80 amino acids** More than 35% identity in the amino acid sequence of the expressed protein (i.e.without the leader sequence, if any), using a window of 80 amino acids and a suitable gap penalty (using Clustal-type alignment programs or equivalent alignment programs). This is one of the recommended test methods of the EFSA scienti c opinion [1], and also of the earlier publication from the FAO/WHO Expert group [2]. The queries were done using Fasta 3.4, using the scripts in appendix C.

2. **35% identity over 80 amino acids (scaled)** Same as item 1, but with scaling enabled. In this way, matches with high identity, but over windows shorter than 80 amino acids can be identified. For example a match with 50% identity over 60 amino acids would still have enough identical amino acids to exceed the 35% threshold over 80 amino acids: $60 \cdot 0.50/80 = 0.375 = 37.5\%$.

3. **Full length alignment** Alignment of glucoamylase from *41SaM2-54*, 41SaM2-54 to each of the allergens, and identify hits with more than 35% identity over the full length of the alignment. These queries were performed using the global alignment "needle", which is an implementation of the Needleman-Wunsch global alignment algorithm [3] in the program package EMBOSS [4].

4. **100% identity over 8 amino acids** Search for identity over 8 contiguous amino acids. The queries were done using Fasta 3.4, using the scripts in appendix C

3 Results

3.1 Database: allergenonline.org

3.1.1 35% identity over 80 amino acids

The following allergens had one or more matches using the method described above

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Schizophyllum commune H4-8	Mushroom	Sch c 1.0101	Food Fungi	Schizophyllum Sch c 1	576	XP_003030591.1	302681819	15

3.1.2 35% identity over 80 amino acids (scaled)

The following allergens had one or more matches using the method described above

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Schizophyllum commune H4-8	Mushroom	Sch c 1.0101	Food Fungi	Schizophyllum Sch c 1	576	XP_003030591.1	302681819	15

3.1.3 Identity over full length

All allergens with more than 10% sequence identity to glucoamylase produced by 41SaM2-54 are shown in appendix D.3. The identities to the allergens identified by the 35% identity over 80 amino acids method (described above) are shown below.

```
P53BAQ_XP_003030591.1 Identity: 294/614  47.9%
```

3.1.4 100% identity over 8 amino acids

```
P53BAQ_window8_103 - 8 aa XP_003030591.1 100.000% identity in 8 aa overlap
P53BAQ_window8_104 - 8 aa XP_003030591.1 100.000% identity in 8 aa overlap
P53BAQ_window8_121 - 8 aa XP_003030591.1 100.000% identity in 8 aa overlap
P53BAQ_window8_122 - 8 aa XP_003030591.1 100.000% identity in 8 aa overlap
P53BAQ_window8_123 - 8 aa XP_003030591.1 100.000% identity in 8 aa overlap
P53BAQ_window8_124 - 8 aa XP_003030591.1 100.000% identity in 8 aa overlap
P53BAQ_window8_125 - 8 aa XP_003030591.1 100.000% identity in 8 aa overlap
P53BAQ_window8_126 - 8 aa XP_003030591.1 100.000% identity in 8 aa overlap
P53BAQ_window8_221 - 8 aa XP_003030591.1 100.000% identity in 8 aa overlap
P53BAQ_window8_222 - 8 aa XP_003030591.1 100.000% identity in 8 aa overlap
```

References

[1] Scientific opinion on the assessment of allergenicity of GM plants and microorganisms and derived food and feed. EFSA panel on genetically Modified Organisms (GMO panel). European Food Safety Authority (EFSA), Parma 2010. (The document may be downloaded from `http://www.efsa.europa.eu/en/scdocs/scdoc/1700.htm`)

[2] Evaluation of Allergenicity of Genetically Modified Foods (Report of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology 22 - 25 January 2001), Food and Agriculture Organization of the United Nations (FAO), Rome 2001. `http://www.who.int/foodsafety/publications/biotech/ec_jan2001/en/`

[3] Needleman, S. B. and Wunsch, C. D. (1970) *J. Mol. Biol.* **48**, p 443-453.

[4] Rice,P. Longden,I. and Bleasby,A. (2000): "EMBOSS: The European Molecular Biology Open Software Suite" *Trends in Genetics* **16**, No 6. p 276-277

A Scripts for toxin homology search

Python scripts for alignment of sequences to the glucoamylase sequence and calculation of sequence lengths and identities. First the script used to run the alignments. The script is stored in a file called "forClustalw.py".

```
def test_clusta_output(clustalw_file, protein, final_target, toxin_name):
    data = clustalw_file
    toxin_sequence = ""
    quary_name = protein.split()[0].replace('>', '')
    star_list = []
    identity = 0
    for line in data:
        if line.startswith("\n"):
            continue
        else:
            temp_list = line.replace("-", "").split()
            if len(temp_list) == 2 and 'CLUSTAL' and temp_list[0] != quary_name not in temp_list:
                if not any('*' in c for c in temp_list):
                    toxin_sequence += temp_list[1]
            else:
                if any('*' in c for c in temp_list):
                    star_list += temp_list

    for i in range(0, len(star_list)):
        for char in star_list[i]:
            if char == '*':
                identity += 1
    final_target.write((toxin_name + "␣" + str(len(toxin_sequence)) + "␣" + str(identity)).replace('\n', '') + "\n")
    data.close()


def make_forClustalw(quary_sequence, toxin_database, final_target, enzym):
    lines = toxin_database.readlines()

    counter0 = 0
    counter1 = 1

    for i in range(0,(len(lines)/2)):
        tmp_file = open('results/clustalw/forClustalw.fasta', 'w')
        tmp_file.write(quary_sequence)
        toxin_name =  lines[counter0][1:]
        tmp_file.write(lines[counter0])
        tmp_file.write(lines[counter1])
        tmp_file.close()

        os.system('/z/linux/bin/clustalw␣' + os.getcwd() + '/results/clustalw/forClustalw.fasta␣>␣/dev/null')

        output = open('results/clustalw/forClustalw.aln')

        test_clusta_output(output, quary_sequence, final_target, toxin_name)

        counter0 += 2
        counter1 += 2
```

% Afterwards the sequence length and identity information can be found in the file summary. This file is processed through the following Python script, which calculates the percentages as described in the text.

B Toxin homology results

UNIPROT entries, that contain the word "toxin", but not "fragment" in the description field and their identity to *41SaM2-54* glucoamylase. The columns are

1. Sequence database accession number
2. Sequence length
3. Number of identical residues after alignment to glucoamylase from 41SaM2-54
4. Percent identity compared to glucoamylase from 41SaM2-54 or the sequence, whichever is longest.
5. Sequence description

Sequences with sequence identity >10% or the 1000 sequences with largest identity are shown.

A0A0H6P4K2 494 96 16.2 RTX toxin OS=Vibrio cholerae OX=666 GN=ERS013186_03642 PE=4 SV=1
I3TK64 524 94 15.9 Rhizobiocin/RTX toxin and hemolysin-type calcium binding protein OS=Tistrella mobilis (strain KA081020-065) O
A0A090P7J4 556 94 15.9 RTX toxin putative OS=Vibrio ponticus OX=265668 GN=JCM19238_3485 PE=4 SV=1
A0A0M9VUF0 582 92 15.6 Putative HC-toxin efflux carrier OS=Escovopsis weberi OX=150374 GN=ESCO_005673 PE=4 SV=1
E4ZFY5 560 91 15.4 Similar to MFS toxin efflux pump (AflT) OS=Leptosphaeria maculans (strain JN3 / isolate v23.1.3 / race Av1-4-
A0A194VDS5 553 91 15.4 Putative HC-toxin efflux carrier TOXA OS=Valsa mali var. pyri OX=694573 GN=VP1G_09081 PE=4 SV=1
A0A0S7DYT1 552 91 15.4 Putative HC-toxin efflux carrier TOXA OS=Aspergillus lentulus OX=293939 GN=ALT_4900 PE=4 SV=1
A0A0M8ZDG6 460 91 15.4 Rtx toxin hemolysin-type calcium-binding protein OS=Asanoa ferruginea OX=53367 GN=ADL14_21925 PE=4 SV=1
A0A239ERZ0 613 94 15.3 Ca2+-binding protein, RTX toxin-related OS=Antarctobacter heliothermus OX=74033 GN=SAMN04488078_101630 PE
A0A2G5I8P7 607 92 15.2 Putative HC-toxin efflux carrier TOXA OS=Cercospora beticola OX=122368 GN=CB0940_00841 PE=4 SV=1
A0A179H4P9 558 90 15.2 Pertussis toxin, subunit 1 domain-containing protein OS=Purpureocillium lilacinum OX=33203 GN=VFPBJ_03261
I9WQY7 455 89 15.1 Rtx toxin hemolysin-type calcium-binding protein OS=Methylobacterium sp. GXF4 OX=1096546 GN=WYO_4383 PE=4 SV=
H0EBY9 572 89 15.1 Hemolysin-type calcium-binding toxin OS=Patulibacter medicamentivorans OX=1097667 GN=PAI11_43740 PE=4 SV=1
F1ZBF7 527 89 15.1 Rhizobiocin/RTX toxin and hemolysin-type calcium binding protein OS=Novosphingobium nitrogenifigens DSM 19370
A0A1H6KML9 566 89 15.1 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Rheinheimera pacifica OX=173990 GN=SAMN05
A0A0A1DNU1 596 90 15.1 Alkaline phosphatase OS=Nocardioides simplex OX=2045 GN=KR76_23415 PE=4 SV=1
S7I480 740 111 15.0 Cytolysin and hemolysin, HlyA, Pore-forming toxin OS=Vibrio fluvialis PG41 OX=1336752 GN=L910_4485 PE=4 SV=1
A0A1W2D0B1 659 99 15.0 Ca2+-binding protein, RTX toxin-related OS=Pseudooceanicola flagellatus OX=1387277 GN=SAMN06295998_11018
A0A1I0YSQ8 749 112 15.0 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Cohnella sp. OV330 OX=1855288 GN=SAMN052
A0A1H0UQP7 714 107 15.0 Ca2+-binding protein, RTX toxin-related OS=Sulfitobacter litoralis OX=335975 GN=SAMN04488512_12738 PE=4
R6S645 363 88 14.9 Putative toxin regulator OS=Lactobacillus ruminis CAG:367 OX=1263085 GN=BN628_00096 PE=4 SV=1
M5B8F6 434 88 14.9 Toxin-antitoxin system, toxin component OS=Clavibacter nebraskensis NCPPB 2581 OX=1097677 GN=CMN_00626 PE=4 S
G3KGV3 472 88 14.9 RTX toxin-related Ca2+-binding protein OS=Adineta vaga OX=104782 PE=4 SV=1
F7R1E6 363 88 14.9 Putative toxin regulator OS=Lactobacillus ruminis SPM0211 OX=1040964 GN=LRU_01515 PE=4 SV=1
A0A2E7YBW9 460 88 14.9 Rtx toxin hemolysin-type calcium-binding protein OS=Methylobacterium sp. OX=409 GN=CMH16_09795 PE=4 SV=1
A0A0M8ZOR1 460 88 14.9 Rtx toxin hemolysin-type calcium-binding protein OS=Streptomyces purpurogeneiscleroticus OX=68259 GN=ADL1
A0A0K8LQ91 572 88 14.9 Putative HC-toxin efflux carrier TOXA OS=Aspergillus udagawae OX=91492 GN=AUD_9430 PE=4 SV=1
A0A0J9X483 582 88 14.9 Similar to Saccharomyces cerevisiae YHR032W ERC1 Member of the multi-drug and toxin extrusion (MATE) fami
A0A0G8G9Y5 363 88 14.9 Putative toxin regulator OS=Lactobacillus ruminis OX=1623 GN=LRP_4 PE=4 SV=1
A0A0F6GGB2 800 119 14.9 Cytolethal distending toxin A/C family protein OS=Burkholderia pseudomallei MSHR2543 OX=1249472 GN=BG16_
A0A034TV62 637 94 14.8 RTX toxin OS=Vibrio sp. JCM 18905 OX=1298600 GN=JCM18905_3690 PE=4 SV=1
N1RZR2 529 87 14.7 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 4) OX=1229665 GN=FOC4
A0A2B1SX95 515 87 14.7 Mosquitocidal toxin protein OS=Bacillus cereus OX=1396 GN=CON36_16980 PE=4 SV=1
A0A1V0G6H9 639 94 14.7 Toxin OS=Neisseria meningitidis OX=487 GN=A6J54_05350 PE=4 SV=1
A0A1L6L1Y0 538 87 14.7 RTX toxin OS=Minicystis rosea OX=888845 GN=A7982_01089 PE=4 SV=1
A0A1L3I282 645 95 14.7 RTX toxin OS=Phaeobacter porticola OX=1844006 GN=PhaeoP97_00770 PE=4 SV=1
A0A1H6QWE8 695 102 14.7 Ca2+-binding protein, RTX toxin-related OS=Cribrihabitans marinus OX=1227549 GN=SAMN05444007_101318 PE=4
A0A194W9S6 553 87 14.7 Putative HC-toxin efflux carrier TOXA OS=Valsa mali OX=105487 GN=VM1G_08824 PE=4 SV=1
A0A0S7DYU0 572 87 14.7 Putative HC-toxin efflux carrier TOXA OS=Aspergillus lentulus OX=293939 GN=ALT_5413 PE=4 SV=1
A0A0F6YL68 447 87 14.7 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Sandaracinus amylolyticus OX=927083
A0A017HRE7 526 87 14.7 RTX toxin OS=Rubellimicrobium mesophilum DSM 19309 OX=442562 GN=Rumeso_02102 PE=4 SV=1
A0A014Q737 435 87 14.7 Toxin HipA OS=Comamonas aquatica DA1877 OX=1457173 GN=AX13_08175 PE=4 SV=1
N4TZB6 646 94 14.6 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 1) OX=1229664 GN=FOC1
I4YL09 525 86 14.6 Ca2+-binding protein, RTX toxin OS=Microvirga lotononidis OX=864069 GN=MicloDRAFT_00053660 PE=4 SV=1
A0A286A7D5 580 86 14.6 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Pedobacter xixiisoli OX=147646
A0A285DP90 537 86 14.6 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Streptomyces microflavus OX=19
A0A242YDF4 515 86 14.6 Mosquitocidal toxin protein OS=Bacillus thuringiensis serovar novosibirsk OX=257986 GN=BK719_08095 PE=4 S
A0A238L6N3 646 94 14.6 Toxin RTX-I translocation ATP-binding protein OS=Pelagimonas varians OX=696760 GN=apxIB_2 PE=4 SV=1
A0A225QRM9 434 86 14.6 RTX toxin OS=Alkalimicrobium pacificum OX=1463601 GN=CDZ97_12525 PE=4 SV=1
A0A1W5D7F9 560 86 14.6 Toxin biosynthesis regulatory protein OS=Umbilicaria pustulata OX=136370 PE=4 SV=1
A0A1L2Z0V8 515 86 14.6 Mosquitocidal toxin protein OS=Bacillus thuringiensis subsp. israelensis OX=1430 GN=ATN07_29875 PE=4 SV=1
A0A1G5K4E2 544 86 14.6 Ca2+-binding protein, RTX toxin-related OS=Microvirga guangxiensis OX=549386 GN=SAMN02927923_03046 PE=4 S
A0A1E3DJZ3 434 86 14.6 RTX toxin OS=Ruegeria sp. PBVC088 OX=1858797 GN=A9320_00470 PE=4 SV=1
A0A194VRK6 543 86 14.6 Putative HC-toxin efflux carrier TOXA OS=Valsa mali OX=105487 GN=VM1G_02315 PE=4 SV=1
A0A167U757 571 86 14.6 Putative HC-toxin efflux carrier TOXA OS=Penicillium chrysogenum OX=5076 GN=EN45_075800 PE=4 SV=1
A0A109JTR7 440 86 14.6 Toxin HipA OS=Rhizobium altiplani OX=1864509 GN=AS026_38615 PE=4 SV=1
A0A0Q9H7P1 515 86 14.6 Mosquitocidal toxin protein OS=Bacillus sp. Root131 OX=1736451 GN=ASE54_26205 PE=4 SV=1
A0A0Q9G6J9 515 86 14.6 Mosquitocidal toxin protein OS=Bacillus sp. Root11 OX=1736425 GN=ASE53_30255 PE=4 SV=1
A0A0N1L0N7 458 86 14.6 Toxin HipA OS=beta proteobacterium AAP51 OX=1523421 GN=D621_15895 PE=4 SV=1
A0A0N1H9Q1 563 86 14.6 Putative HC-toxin efflux carrier TOXA OS=Phialophora attae OX=1664694 GN=AB675_4678 PE=4 SV=1
A0A0F6SHR9 505 86 14.6 RTX toxin OS=Sandaracinus amylolyticus OX=927083 GN=DB32_008073 PE=4 SV=1
A0A0B3S168 434 86 14.6 RTX toxin OS=Mameliella alba OX=561184 GN=OA50_02744 PE=4 SV=1
A0A086Y4T8 440 86 14.6 Rtx toxin hemolysin-type calcium-binding protein OS=Haematobacter massiliensis OX=195105 GN=CN97_16600 PE
A0A074T8D9 637 93 14.6 RTX toxin OS=Thioclava dalianensis OX=1185766 GN=DL1_16855 PE=4 SV=1
B5X306 635 92 14.5 Multidrug and toxin extrusion protein OS=Salmo salar OX=8030 GN=S47A1 PE=2 SV=1
A0A2K7RL94 640 93 14.5 Toxin OS=Neisseria meningitidis OX=487 GN=A6J49_10025 PE=4 SV=1
A0A2A9D2X0 691 100 14.5 Restriction endonuclease fold toxin 5 of polymorphic toxin system OS=Serinibacter salmoneus OX=556530 GN
A0A0L6CYW6 684 99 14.5 RTX-I toxin determinant A from serotypes 1/9 OS=Roseovarius tolerans OX=74031 GN=apxIA PE=4 SV=1

Q489Z5 455 85 14.4 Zona occludens toxin OS=Colwellia psychrerythraea (strain 34H / ATCC BAA-681) OX=167879 GN=CPS_0361 PE=4 SV=1
Q487Y5 455 85 14.4 Zona occludens toxin-like protein OS=Colwellia psychrerythraea (strain 34H / ATCC BAA-681) OX=167879 GN=CPS_0
Q0FTY9 521 85 14.4 Toxin secretion ABC transporter protein, HlyB family OS=Pelagibaca bermudensis (strain JCM 13377 / KCTC 12554
M5U053 525 85 14.4 Toxin secretion ABC transporter ATP-binding protein OS=Rhodopirellula sallentina SM41 OX=1263870 GN=RSSM_0370
A0A2I7HTH9 461 85 14.4 RTX toxin OS=Phaeobacter inhibens OX=221822 GN=PhaeoP54_00239 PE=4 SV=1
A0A2C1S163 512 85 14.4 Mosquitocidal toxin protein OS=Bacillus cereus OX=1396 GN=CN271_19570 PE=4 SV=1
A0A2A9GI27 758 109 14.4 Ca2+-binding RTX toxin-like protein OS=Thioclava sp. ES.031 OX=1798203 GN=AXZ77_1963 PE=4 SV=1
A0A257C8I0 464 85 14.4 Rtx toxin hemolysin-type calcium-binding protein OS=Burkholderiales bacterium PBB6 OX=2015568 GN=CFE46_18
A0A243K2D3 515 85 14.4 Mosquitocidal toxin protein OS=Bacillus thuringiensis serovar argentinensis OX=180880 GN=BK740_18710 PE=4
A0A1Z1CCX4 573 85 14.4 Putative HC-toxin efflux carrier OS=Cladonia uncialis subsp. uncialis OX=180999 PE=4 SV=1
A0A1Y5THU4 646 93 14.4 Toxin RTX-I translocation ATP-binding protein OS=Tropicibacter litoreus R37 OX=1200284 GN=apxIB_2 PE=4 SV
A0A1W2TAN0 477 85 14.4 Putative toxin biosynthesis protein OS=Rosellinia necatrix OX=77044 GN=SAMD00023353_3001240 PE=4 SV=1
A0A1S9TJ03 515 85 14.4 Mosquitocidal toxin protein OS=Bacillus cereus OX=1396 GN=BW897_24925 PE=4 SV=1
A0A1S8TMJ7 507 85 14.4 Toxin B OS=Clostridium sp. BL-8 OX=349938 GN=toxB_2 PE=4 SV=1
A0A1L7R8H8 360 85 14.4 Putative toxin 43 OS=Actinomyces succiniciruminis OX=1522002 GN=AAM4_0172 PE=4 SV=1
A0A1L6LRJ9 456 85 14.4 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Minicystis rosea OX=888845 GN=A7982_
A0A1G5X2X7 457 85 14.4 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Pseudomonas sp. NFPP33 OX=1566
A0A1C6FCL1 556 85 14.4 Toxin A OS=uncultured Ruminococcus sp. OX=165186 GN=toxA_2 PE=4 SV=1
A0A1C5PA52 556 85 14.4 Toxin A OS=uncultured Ruminococcus sp. OX=165186 GN=toxA_3 PE=4 SV=1
A0A1C5KJG5 460 85 14.4 Toxin A OS=uncultured Clostridium sp. OX=59620 GN=toxA_2 PE=4 SV=1
A0A194VBE4 599 86 14.4 Putative HC-toxin efflux carrier TOXA OS=Valsa mali var. pyri OX=694573 GN=VP1G_08391 PE=4 SV=1
A0A167KLX3 580 85 14.4 MFS toxin efflux pump (AflT) OS=Metarhizium rileyi RCEF 4871 OX=1081105 GN=NOR_00489 PE=4 SV=1
A0A0W0YK51 790 114 14.4 Structural toxin protein RtxA OS=Legionella shakespearei DSM 23087 OX=1122169 GN=rtxA-1_2 PE=4 SV=1
A0A0U1LR64 546 85 14.4 Putative HC-toxin efflux carrier TOXA OS=Talaromyces islandicus OX=28573 GN=PISL3812_02613 PE=4 SV=1
A0A0L0MHF3 461 85 14.4 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Candidatus Burkholderia verschuereni
A0A0C4YET3 428 85 14.4 Multidrug and toxin extrusion (MATE) family efflux pump, YdhE/NorM-like OS=Cupriavidus basilensis OX=6889
U6NP82 651 93 14.3 Proteinase inhibitor and Peptidase M14 and Metridin ShK toxin domain containing protein OS=Haemonchus contort
A0A1Y6BXN2 645 92 14.3 Pre-toxin TG OS=Pseudobacteriovorax antillogorgiicola OX=1513793 GN=SAMN06296036_109179 PE=4 SV=1
A0A1P8V0N6 677 97 14.3 Ca2+-binding protein, RTX toxin OS=Pelagibaca abyssi OX=1250539 GN=Ga0080574_TMP4931 PE=4 SV=1
A0A1I4LYB1 698 100 14.3 Ca2+-binding protein, RTX toxin-related OS=Methylobacterium salsuginis OX=414703 GN=SAMN04488125_1337 PE
A0A1H2S1U3 706 101 14.3 Ca2+-binding protein, RTX toxin-related OS=Ruegeria mobilis OX=379347 GN=SAMN05444385_101648 PE=4 SV=1
A0A1H2C7E9 677 97 14.3 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Pseudomonas chlororaphis OX=587753 GN=SAM
A0A0J8VED3 610 87 14.3 Cytolysin and hemolysin HlyA Pore-forming toxin OS=Photobacterium swingsii OX=680026 GN=AB733_07730 PE=4
A0A085TUR7 623 89 14.3 RTX toxin OS=Thioclava atlantica OX=1317124 GN=DW2_12965 PE=4 SV=1
S5BVR7 606 86 14.2 RTX toxin OS=Alteromonas mediterranea UM7 OX=1300258 GN=I635_04940 PE=4 SV=1
R9PK43 489 84 14.2 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Agarivorans albus MKT 106 OX=1331007 GN=
P38179 458 84 14.2 Dol-P-Man:Man(5)GlcNAc(2)-PP-Dol alpha-1,3-mannosyltransferase OS=Saccharomyces cerevisiae (strain ATCC 20450
N4TPX1 565 84 14.2 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 1) OX=1229664 GN=FOC1
L1KRP1 505 84 14.2 Zeta toxin OS=Streptomyces ipomoeae 91-03 OX=698759 GN=STRIP9103_06495 PE=4 SV=1
A0A2H3KY63 515 84 14.2 Mosquitocidal toxin protein OS=Bacillus toyonensis OX=155322 GN=CON90_19740 PE=4 SV=1
A0A2B7A003 517 84 14.2 Mosquitocidal toxin protein OS=Bacillus thuringiensis OX=1428 GN=COF70_06935 PE=4 SV=1
A0A2A8KEA9 517 84 14.2 Mosquitocidal toxin protein OS=Bacillus cereus OX=1396 GN=CN508_08710 PE=4 SV=1
A0A292ZJV5 483 84 14.2 Multidrug and toxin extrusion family efflux pump YdhE OS=Sphingobium fuliginis (strain ATCC 27551) OX=336
A0A267FVZ4 627 89 14.2 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig030132g2 PE=3 SV=1
A0A267EN60 627 89 14.2 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig030132g1 PE=3 SV=1
A0A243MZA1 515 84 14.2 Mosquitocidal toxin protein OS=Bacillus thuringiensis serovar sinensis OX=1042875 GN=BK788_15525 PE=4 SV=
A0A242WRN3 515 84 14.2 Mosquitocidal toxin protein OS=Bacillus thuringiensis serovar cameroun OX=180885 GN=BK702_11655 PE=4 SV=1
A0A238HCN5 461 84 14.2 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Burkholderia singularis OX=1503053 G
A0A226MH28 534 84 14.2 Multidrug and toxin extrusion protein OS=Callipepla squamata OX=9009 GN=ASZ78_002883 PE=3 SV=1
A0A1S8X4F5 552 84 14.2 Multidrug and toxin extrusion protein OS=Opisthorchis viverrini OX=6198 GN=X801_02498 PE=3 SV=1
A0A1S1VAW5 555 84 14.2 MFS toxin efflux pump OS=Colletotrichum incanum OX=1573173 GN=CSPAE12_11910 PE=4 SV=1
A0A1N6YX31 488 84 14.2 Aerolysin/Pertussis toxin (APT) domain-containing protein OS=Aeromonas veronii OX=654 GN=SAMN05892873_130
A0A1L6LGT0 438 84 14.2 RTX toxin OS=Minicystis rosea OX=888845 GN=A7982_06270 PE=4 SV=1
A0A1L6LGJ3 520 84 14.2 RTX toxin OS=Minicystis rosea OX=888845 GN=A7982_06179 PE=4 SV=1
A0A1I5VLA6 524 84 14.2 Ca2+-binding protein, RTX toxin-related OS=Enterovibrio norvegicus DSM 15893 OX=1121869 GN=SAMN03084138_0
A0A1I3MXE0 599 85 14.2 Novel toxin 15 OS=Jannaschia pohangensis OX=390807 GN=SAMN04488095_1988 PE=4 SV=1
A0A1I1MTF1 486 84 14.2 Aerolysin/Pertussis toxin (APT) domain-containing protein OS=Pseudoalteromonas denitrificans DSM 6059 OX=
A0A1H3CCW0 497 84 14.2 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Thiocapsa roseopersicina OX=10
A0A1C5SKV4 503 84 14.2 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_3 PE=4 SV=1
A0A1A9PZY2 515 84 14.2 Mosquitocidal toxin protein OS=Bacillus wiedmannii OX=1890302 GN=A6280_25240 PE=4 SV=1
A0A194UPP6 506 84 14.2 Putative HC-toxin efflux carrier TOXA OS=Valsa mali var. pyri OX=694573 GN=VP1G_01041 PE=4 SV=1
A0A161XYW4 492 84 14.2 Killer toxin subunits alpha/beta OS=Penicillium chrysogenum OX=5076 GN=EN45_046790 PE=4 SV=1
A0A0S4PCU7 496 84 14.2 Putative RTX toxin hemolysin-type calcium-binding protein OS=Janthinobacterium sp. CG23_2 OX=1706231 GN=B
A0A0N0X719 506 84 14.2 RTX toxin OS=Pseudomonas syringae pv. cilantro OX=81035 GN=ABJ99_0997 PE=4 SV=1
A0A0L0N509 579 84 14.2 Putative HC-toxin efflux carrier TOXA OS=Tolypocladium ophioglossoides CBS 100239 OX=1163406 GN=TOPH_0610
A0A0K8LGY1 541 84 14.2 Putative HC-toxin efflux carrier TOXA OS=Aspergillus udagawae OX=91492 GN=AUD_6536 PE=4 SV=1
A0A0K1Q1I7 538 84 14.2 RTX toxin OS=Labilithrix luteola OX=1391654 GN=AKJ09_06319 PE=4 SV=1
A0A0A1HZB0 467 84 14.2 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Pseudomonas sp. SHC52 OX=984195 GN=B

A0A090EIH4 382 84 14.2 Putative RTX toxin hemolysin-type protein OS=Mesorhizobium plurifarium OX=69974 GN=MPL3356_80002 PE=4 SV=
A0A034TWE2 597 85 14.2 RTX toxin and related Ca2+-binding protein OS=Vibrio sp. JCM 18905 OX=1298600 GN=JCM18905_2594 PE=4 SV=1
A6E153 844 119 14.1 Rhizobiocin/RTX toxin OS=Roseovarius sp. TM1035 OX=391613 GN=RTM1035_19751 PE=4 SV=1
A0A1I7M5W9 708 100 14.1 Ca2+-binding protein, RTX toxin-related OS=Massilia namucuonensis OX=1035707 GN=SAMN05216552_10635 PE=4
A0A0LON1R8 695 98 14.1 Killer toxin subunits alpha/beta OS=Tolypocladium ophioglossoides CBS 100239 OX=1163406 GN=TOPH_07630 PE=
W7WL96 721 101 14.0 RTX-I toxin determinant B OS=Methylibium sp. T29-B OX=1437443 GN=apxIB_1 PE=4 SV=1
W7W2B9 721 101 14.0 RTX-I toxin determinant B OS=Methylibium sp. T29 OX=1430884 GN=apxIB_1 PE=4 SV=1
H2M6N8 559 83 14.0 Anthrax toxin receptor 1 OS=Oryzias latipes OX=8090 GN=ANTXR1 PE=4 SV=1
H0SGY7 582 83 14.0 Putative secretion ATP-binding protein (ABC-type transporter family) toxin/protease secretion system OS=Brady
F2J697 444 83 14.0 Putative toxin/protease secretion system OS=Polymorphum gilvum (strain LMG 25793 / CGMCC 1.9160 / SL003B-26A1
F2G314 606 85 14.0 RTX toxin OS=Alteromonas mediterranea (strain DSM 17117 / CIP 110805 / LMG 28347 / Deep ecotype) OX=1774373 G
C9P1Q9 663 93 14.0 General secretion pathway protein D (Cholera toxin secretion protein epsD) OS=Vibrio metschnikovii CIP 69.14
A4YUZ6 584 83 14.0 Putative secretion ATP-binding protein (ABC-type transporter family) putative toxin/protease secretion system
A0A2K3ILV4 450 83 14.0 Toxin HipA OS=Burkholderia pseudomallei OX=28450 GN=CF641_12765 PE=4 SV=1
A0A2K3IC69 450 83 14.0 Toxin HipA OS=Burkholderia sp. 137 OX=2020483 GN=CF648_11755 PE=4 SV=1
A0A2K3HLD7 450 83 14.0 Toxin HipA OS=Burkholderia sp. 117 OX=2020481 GN=CF647_11800 PE=4 SV=1
A0A2K3GF44 450 83 14.0 Toxin HipA OS=Burkholderia sp. 129 OX=2020482 GN=CF650_10555 PE=4 SV=1
A0A2K3FG61 450 83 14.0 Toxin HipA OS=Burkholderia sp. 136(2017) OX=2020484 GN=CF649_11755 PE=4 SV=1
A0A2K3EM91 450 83 14.0 Toxin HipA OS=Burkholderia pseudomallei OX=28450 GN=CF640_16910 PE=4 SV=1
A0A2J9KSE0 427 83 14.0 Type II toxin-antitoxin system HipA family toxin OS=Mobiluncus mulieris OX=2052 GN=CEP82_010625 PE=4 SV=1
A0A2G5HPV9 599 84 14.0 Putative HC-toxin efflux carrier TOXA OS=Cercospora beticola OX=122368 GN=CB0940_08687 PE=4 SV=1
A0A2E2LW11 485 83 14.0 Aerolysin family beta-barrel pore-forming toxin OS=Pseudoalteromonas sp. OX=53249 GN=CMK69_03170 PE=4 SV=
A0A2B5VLK6 515 83 14.0 Mosquitocidal toxin protein OS=Bacillus toyonensis OX=155322 GN=COL90_28400 PE=4 SV=1
A0A2B5DGZ3 515 83 14.0 Mosquitocidal toxin protein OS=Bacillus toyonensis OX=155322 GN=CN594_20240 PE=4 SV=1
A0A290YG21 450 83 14.0 Toxin HipA OS=Burkholderia mallei OX=13373 GN=RY28_18575 PE=4 SV=1
A0A267FY51 627 88 14.0 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig030132g3 PE=3 SV=1
A0A1W6C0H4 504 83 14.0 Zeta toxin family protein OS=Pseudomonas fluorescens OX=294 PE=4 SV=1
A0A1V5QBB1 552 83 14.0 Toxin A OS=Firmicutes bacterium ADurb.Bin354 OX=1852889 GN=toxA PE=4 SV=1
A0A1V4J174 756 106 14.0 Toxin B OS=Clostridium chromiireducens OX=225345 GN=toxB PE=4 SV=1
A0A1S8B9G0 488 83 14.0 Killer toxin subunits alpha/beta OS=Diplodia seriata OX=420778 GN=BK809_0000134 PE=3 SV=1
A0A1N7KEV1 528 83 14.0 Ca2+-binding protein, RTX toxin-related OS=Gemmobacter megaterium OX=1086013 GN=SAMN05421774_101385 PE=4
A0A1I6XUJ3 485 83 14.0 Aerolysin/Pertussis toxin (APT) domain-containing protein OS=Pseudoalteromonas sp. DSM 26666 OX=1761892 G
A0A1H7ZAT8 788 110 14.0 Ca2+-binding protein, RTX toxin-related OS=Gemmobacter aquatilis OX=933059 GN=SAMN04488103_101438 PE=4 S
A0A1C5SFE3 498 83 14.0 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_4 PE=4 SV=1
A0A1C5LG66 457 83 14.0 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_1 PE=4 SV=1
A0A1C5KHQ2 599 84 14.0 Toxin A OS=uncultured Clostridium sp. OX=59620 GN=toxA_1 PE=4 SV=1
A0A1C1CND1 573 83 14.0 Putative HC-toxin efflux carrier TOXA OS=Cladophialophora carrionii OX=86049 GN=TOXA PE=4 SV=1
A0A194VAH5 569 83 14.0 Putative HC-toxin efflux carrier TOXA OS=Valsa mali var. pyri OX=694573 GN=VP1G_08084 PE=4 SV=1
A0A105TH13 719 101 14.0 RTX-I toxin determinant B OS=Pseudomonas sp. TAD18 OX=1729583 GN=apxIB PE=4 SV=1
A0A0R2AQP6 370 83 14.0 Toxin regulator OS=Lactobacillus agilis DSM 20509 OX=1423718 GN=FC14_GL001994 PE=4 SV=1
A0A0G2K7L7 608 85 14.0 Anthrax toxin receptor-like OS=Rattus norvegicus OX=10116 GN=Antxrl PE=4 SV=1
A0A0G2JSJ5 576 83 14.0 Anthrax toxin receptor 1 OS=Rattus norvegicus OX=10116 GN=Antxr1 PE=1 SV=1
A0A0E9DZR3 451 83 14.0 Toxin A OS=Chlamydia trachomatis OX=813 GN=toxA_2 PE=4 SV=1
A0A090R2V1 475 83 14.0 RTX toxin putative OS=Vibrio sp. C7 OX=1001886 GN=JCM19233_772 PE=4 SV=1
A0A086T2G4 573 83 14.0 Putative HC-toxin efflux carrier-like protein OS=Acremonium chrysogenum (strain ATCC 11550 / CBS 779.69 /
W7J2K5 401 82 13.9 RTX toxin and related Ca2+-binding protein OS=Actinokineospora spheciospongiae OX=909613 GN=UO65_1511 PE=4 SV
V4B2N0 520 82 13.9 Multidrug and toxin extrusion protein OS=Lottia gigantea OX=225164 GN=LOTGIDRAFT_157993 PE=3 SV=1
R6HLG6 389 82 13.9 Predicted membrane protein putative toxin regulator OS=Firmicutes bacterium CAG:552 OX=1263029 GN=BN704_00985
Q2K6Y0 538 82 13.9 Putative RTX toxin hemolysin-type protein OS=Rhizobium etli (strain CFN 42 / ATCC 51251) OX=347834 GN=RHE_CH0
Q15Z74 559 82 13.9 Toxin secretion ABC transporter ATP-binding protein OS=Pseudoalteromonas atlantica (strain T6c / ATCC BAA-108
N4WEA4 475 82 13.9 3-hydroxyacyl-CoA dehydrogenase-like protein LAM1 OS=Cochliobolus heterostrophus (strain C4 / ATCC 48331 / ra
K7A6Q3 407 82 13.9 RTX toxin, putative OS=Paraglaciecola polaris LMG 21857 OX=1129793 GN=GPLA_0216 PE=4 SV=1
A0A2G7FLE8 531 82 13.9 MFS toxin efflux pump OS=Aspergillus arachidicola OX=656916 GN=AARAC_009776 PE=4 SV=1
A0A2G5L143 435 82 13.9 Toxin HipA OS=Reichenbachiella sp. 5M10 OX=1889772 GN=BFP72_06040 PE=4 SV=1
A0A2G5GQA3 597 83 13.9 Putative HC-toxin efflux carrier TOXA OS=Cercospora beticola OX=122368 GN=CB0940_12234 PE=4 SV=1
A0A2G1BAX8 485 82 13.9 Aerolysin family beta-barrel pore-forming toxin OS=Pseudoalteromonas sp. 3D05 OX=2045444 GN=CSC79_02665 P
A0A2E6YND9 401 82 13.9 Addiction module toxin YoeB OS=Halieaceae bacterium OX=2026743 GN=CME58_02695 PE=4 SV=1
A0A2C4X2E3 515 82 13.9 Mosquitocidal toxin protein OS=Bacillus toyonensis OX=155322 GN=COM61_26895 PE=4 SV=1
A0A292C8N8 457 82 13.9 Toxin HipA OS=Proteus vulgaris OX=585 GN=CRN77_06665 PE=4 SV=1
A0A259FEJ6 436 82 13.9 Toxin HipA OS=Hydrogenophilales bacterium 17-64-34 OX=1970527 GN=B7X87_12990 PE=4 SV=1
A0A226NWM4 534 82 13.9 Multidrug and toxin extrusion protein OS=Colinus virginianus OX=9014 GN=H355_004230 PE=3 SV=1
A0A212S3T8 697 97 13.9 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Arboriscoccus pini OX=1963835
A0A1V8QCT9 463 82 13.9 Papain fold toxin 1 OS=Bifidobacterium adolescentis OX=1680 GN=B5789_0693 PE=4 SV=1
A0A1V0PZJ9 454 82 13.9 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Rhodovulum sp. P5 OX=1564506 GN=RGUI
A0A1S8TLH9 492 82 13.9 Toxin A OS=Clostridium sp. BL-8 OX=349938 GN=toxA_1 PE=4 SV=1
A0A1S8BIL0 604 84 13.9 Putative HC-toxin efflux carrier TOXA OS=Diplodia seriata OX=420778 GN=BK809_0007214 PE=4 SV=1
A0A1Q9QFI7 433 82 13.9 RTX toxin OS=Rhodovulum sulfidophilum OX=35806 GN=BV509_15130 PE=4 SV=1
A0A1Q8RLU3 570 82 13.9 Putative HC-toxin efflux carrier TOXA 14 OS=Colletotrichum chlorophyti OX=708187 GN=CCHL11_04350 PE=4 SV=

A0A1N7PI78 728 101 13.9 Ca2+-binding protein, RTX toxin-related OS=Roseivivax lentus OX=633194 GN=SAMN05421759_1163 PE=4 SV=1
A0A1N6TZ39 656 91 13.9 Ca2+-binding protein, RTX toxin-related OS=Azoarcus tolulyticus OX=34027 GN=SAMN05421829_105162 PE=4 SV=1
A0A1M3MMP1 435 82 13.9 Toxin HipA OS=Rhizobiales bacterium 63-22 OX=1895812 GN=BGP07_01150 PE=4 SV=1
A0A1I3FGK7 554 82 13.9 Ca2+-binding protein, RTX toxin-related OS=Paracoccus aminovorans OX=34004 GN=SAMN04488021_1742 PE=4 SV=1
A0A1I2JE42 484 82 13.9 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Thermoflexibacter ruber OX=1003 GN=SAMN04
A0A1I1PRS7 768 107 13.9 Ca2+-binding protein, RTX toxin-related OS=Devosia psychrophila OX=728005 GN=SAMN04488059_12232 PE=4 SV=
A0A1H3WRC7 599 83 13.9 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Prevotella sp. tc2-28 OX=17618
A0A1E5CBC5 588 82 13.9 RTX toxin OS=Enterovibrio norvegicus FF-454 OX=1185651 GN=A1OK_19245 PE=4 SV=1
A0A1C6JKE0 457 82 13.9 Toxin A OS=uncultured Clostridium sp. OX=59620 GN=toxA_1 PE=4 SV=1
A0A1C6I9D9 457 82 13.9 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_3 PE=4 SV=1
A0A1C6ASN1 457 82 13.9 Toxin A OS=uncultured Ruminococcus sp. OX=165186 GN=toxA_1 PE=4 SV=1
A0A1C5X7J4 498 82 13.9 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_4 PE=4 SV=1
A0A1C5VY57 496 82 13.9 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_1 PE=4 SV=1
A0A1C4YR25 523 82 13.9 Nucleotidyl transferase AbiEii toxin, Type IV TA system OS=Micromonospora chokoriensis OX=356851 GN=GA007
A0A194VSD4 569 82 13.9 Putative HC-toxin efflux carrier TOXA OS=Valsa mali OX=105487 GN=VM1G_01882 PE=4 SV=1
A0A174T5L5 490 82 13.9 Toxin A OS=Blautia obeum OX=40520 GN=toxA_3 PE=4 SV=1
A0A174NC28 490 82 13.9 Toxin A OS=Blautia wexlerae OX=418240 GN=toxA_2 PE=4 SV=1
A0A174DNZ5 490 82 13.9 Toxin A OS=Blautia wexlerae OX=418240 GN=toxA_2 PE=4 SV=1
A0A174D9G6 490 82 13.9 Toxin A OS=Blautia obeum OX=40520 GN=toxA_3 PE=4 SV=1
A0A173WZ40 490 82 13.9 Toxin A OS=Blautia obeum OX=40520 GN=toxA_1 PE=4 SV=1
A0A0W1A043 578 82 13.9 Structural toxin protein RtxA OS=Legionella waltersii OX=66969 GN=rtxA_1 PE=4 SV=1
A0A0U1LKL9 552 82 13.9 Putative HC-toxin efflux carrier TOXA OS=Talaromyces islandicus OX=28573 GN=PISL3812_00846 PE=4 SV=1
A0A0K3CGB7 468 82 13.9 BY PROTMAP: gi|472583216|gb|EMS20870.1| toxin biosynthesis protein [Rhodosporidium toruloides NP11] gi|64
A0A0H6RD74 549 82 13.9 RTX toxin OS=Vibrio cholerae OX=666 GN=ERS013200_01027 PE=4 SV=1
A0A0D0QK68 507 82 13.9 RTX toxin OS=Wenxinia marina DSM 24838 OX=1123501 GN=Wenmar_04113 PE=4 SV=1
A0A0C2ETC3 400 82 13.9 RTX toxin, Ca2+-binding protein OS=Pseudomonas batumici OX=226910 GN=UCMB321_4446 PE=4 SV=1
A0A0A2VQG7 577 82 13.9 Putative HC-toxin efflux carrier TOXA OS=Beauveria bassiana D1-5 OX=1245745 GN=BBAD15_g6230 PE=4 SV=1
A0A0A0XLM5 656 91 13.9 Type III secretion toxin, effector OS=Bordetella pertussis B1920 OX=743278 GN=bteA PE=4 SV=1
A0A093VE13 572 82 13.9 Putative HC-toxin efflux carrier TOXA OS=Talaromyces marneffei PM1 OX=1077442 GN=GQ26_0071430 PE=4 SV=1
A0A086T7M2 531 82 13.9 Putative HC-toxin efflux carrier-like protein OS=Acremonium chrysogenum (strain ATCC 11550 / CBS 779.69 /
A0A074WYD3 575 82 13.9 MFS toxin efflux pump OS=Aureobasidium namibiae CBS 147.97 OX=1043004 GN=M436DRAFT_78012 PE=4 SV=1
A0A034TXF2 605 84 13.9 RTX toxin and related Ca2+-binding protein OS=Vibrio sp. JCM 18905 OX=1298600 GN=JCM18905_2597 PE=4 SV=1
A0A011JQ08 719 100 13.9 Pre-toxin domain with VENN motif family protein OS=Acinetobacter sp. 723929 OX=1310711 GN=J610_3593 PE=4
U3PTD6 607 84 13.8 Multidrug and toxin extrusion protein OS=Danio rerio OX=7955 GN=slc47a4 PE=2 SV=1
Q7Z2W4-3 624 86 13.8 Isoform 3 of Zinc finger CCCH-type antiviral protein 1 OS=Homo sapiens OX=9606 GN=ZC3HAV1
J1T2A7 782 108 13.8 Ca2+-binding protein, RTX toxin OS=Rhizobium sp. CF142 OX=1144314 GN=PMI11_00049 PE=4 SV=1
E7FEQ7 607 84 13.8 Multidrug and toxin extrusion protein OS=Danio rerio OX=7955 GN=slc47a4 PE=3 SV=1
A0A2H3FQ77 600 83 13.8 Multidrug and toxin extrusion protein OS=Diplocarpon rosae OX=946125 GN=BUE80_DR003486 PE=4 SV=1
A0A1M6NVV0 602 83 13.8 Toxin 50 OS=Clostridium cavendishii DSM 21758 OX=1121302 GN=SAMN02745163_03034 PE=4 SV=1
A0A1M5BJN3 668 92 13.8 Ca2+-binding protein, RTX toxin-related OS=Litoreibacter ascidiaceicola OX=1486859 GN=SAMN05444273_10614
A0A0W0WSS9 615 85 13.8 Structural toxin protein RtxA OS=Legionella moravica OX=39962 GN=rtxA PE=4 SV=1
V3ZGJ0 531 81 13.7 Multidrug and toxin extrusion protein OS=Lottia gigantea OX=225164 GN=LOTGIDRAFT_169500 PE=3 SV=1
Q6DE00 478 81 13.7 2-(3-amino-3-carboxypropyl)histidine synthase subunit 2 OS=Xenopus laevis OX=8355 GN=dph2 PE=2 SV=1
N4TZW0 581 81 13.7 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 1) OX=1229664 GN=FOC1
N4TZE4 529 81 13.7 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 1) OX=1229664 GN=FOC1
N1SO57 559 81 13.7 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 4) OX=1229665 GN=FOC4
M6Q4U2 431 81 13.7 Toxin HINT domain protein OS=Leptospira weilii str. UI 13098 OX=1088542 GN=LEP1GSC108_3379 PE=4 SV=1
H1VBG3 619 85 13.7 Cercosporin toxin biosynthesis protein OS=Colletotrichum higginsianum (strain IMI 349063) OX=759273 GN=CH063_
H0TGV9 652 89 13.7 Toxin secretion ABC transporter (ATP-binding and membrane protein) hlyB-like protein OS=Bradyrhizobium sp. ST
H0TA22 584 81 13.7 Putative secretion ATP-binding protein (ABC-type transporter family) toxin/protease secretion system OS=Brady
G8TEV8 462 81 13.7 Binary exotoxin B/Anthrax toxin B moiety protective antigen OS=Niastella koreensis (strain DSM 17620 / KACC 1
G8S377 487 81 13.7 Hansenula MRAKII killer toxin-resistant protein 1 OS=Actinoplanes sp. (strain ATCC 31044 / CBS 674.73 / SE50/
G3QBK0 577 81 13.7 Multidrug and toxin extrusion protein OS=Gasterosteus aculeatus OX=69293 PE=3 SV=1
D5H918 764 105 13.7 Probable toxin secretion ABC transporter ATP-binding protein, putative OS=Salinibacter ruber (strain M8) OX=
D4ZBF5 746 102 13.7 Toxin secretion ABC transporter protein, HlyB family OS=Shewanella violacea (strain JCM 10179 / CIP 106290 /
C4NZQ8 665 91 13.7 Cry4 delta-toxin-like protein OS=Bacillus thuringiensis OX=1428 PE=3 SV=2
A7KAU2 568 81 13.7 Multidrug and toxin extrusion protein 1 OS=Oryctolagus cuniculus OX=9986 GN=SLC47A1 PE=2 SV=1
A0A2I7M243 523 81 13.7 RTX toxin OS=Phaeobacter inhibens OX=221822 GN=PhaeoP66_03111 PE=4 SV=1
A0A2I6LTH6 707 97 13.7 RTX toxin OS=Acinetobacter pittii OX=48296 GN=BVD86_04345 PE=4 SV=1
A0A2G9BF62 443 81 13.7 Type II toxin-antitoxin system HipA family toxin OS=Escherichia coli OX=562 GN=CT148_13390 PE=4 SV=1
A0A2A9LUG6 565 81 13.7 Antitoxin Phd_YefM of type II toxin-antitoxin system OS=Burkholderia sp. JKS000303 OX=1938747 GN=BX604_32
A0A2A3IKH8 418 81 13.7 Cytolethal distending toxin A/C domain-containing protein OS=Streptomyces sp. 2321.6 OX=1938840 GN=BX261_
A0A2A3HU73 428 81 13.7 Nucleic acid/nucleotide deaminase of polymorphic system toxin OS=Streptomyces sp. TLI_235 OX=1938860 GN=B
A0A267ETL6 597 82 13.7 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig015591g1 PE=3 SV=1
A0A267DC31 597 82 13.7 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig015591g3 PE=3 SV=1
A0A212TWE9 418 81 13.7 Cytolethal distending toxin A/C domain-containing protein OS=Streptomyces sp. 2114.4 OX=1938836 GN=SAMN06
A0A212TGA0 485 81 13.7 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Hymenobacter gelipurpurascens OX=89968 GN
A0A1Y6D2F5 456 81 13.7 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Methylomagnum ishizawai OX=176

A0A1X9PXX4 551 81 13.7 Toxin efflux transporter MFS OS=Phyllosticta cirsii OX=1986016 GN=phyL3 PE=4 SV=1
A0A1X4JRQ1 450 81 13.7 Toxin HipA OS=Burkholderia pseudomallei OX=28450 GN=BOC41_26325 PE=4 SV=1
A0A1W6FVM3 450 81 13.7 Toxin HipA OS=Burkholderia pseudomallei OX=28450 GN=BOC42_35785 PE=4 SV=1
A0A1T5C3I0 510 81 13.7 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Parabacteroides chartae OX=103
A0A1T2B0G3 434 81 13.7 RTX toxin OS=Thioclava sp. DLFJ4-1 OX=1915313 GN=BMI85_10550 PE=4 SV=1
A0A1R4F8U9 576 81 13.7 Putative RTX toxin OS=Actinomycetales bacterium JB111 OX=1434822 GN=CZ771_05400 PE=4 SV=1
A0A1Q8YHB5 447 81 13.7 Toxin hipA OS=Rhodoferax antarcticus ANT.BR OX=1111071 GN=BLL52_1281 PE=4 SV=1
A0A1Q8RX17 526 81 13.7 Putative HC-toxin efflux carrier TOXA 26 OS=Colletotrichum chlorophyti OX=708187 GN=CCHL11_09492 PE=4 SV=
A0A1K2EG61 488 81 13.7 Pre-toxin domain with VENN motif-containing protein OS=Enterobacter sp. NFIX03 OX=1566257 GN=SAMN03159436
A0A1J5R692 724 99 13.7 Toxin RTX-I translocation ATP-binding protein OS=mine drainage metagenome OX=410659 GN=apxIB_7 PE=4 SV=1
A0A1J1KLC2 663 91 13.7 Putative haemolysin-type calcium-binding toxin, RTX-like (Expressed) OS=Planktothrix serta PCC 8927 OX=67
A0A1H8XEL1 737 101 13.7 Pre-toxin domain with VENN motif-containing protein OS=Acinetobacter sp. UNC434CL69Tsu2S25 OX=1502768 GN
A0A1H6H5Q5 553 81 13.7 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Rhizobium sp. NFR12 OX=1566261
A0A1H2CXW5 474 81 13.7 Ca2+-binding protein, RTX toxin-related OS=Actinoplanes derwentensis OX=113562 GN=SAMN04489716_7228 PE=4
A0A1G6WST4 626 86 13.7 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Myxococcus virescens OX=83456 GN=SAMN0448
A0A1E3KY31 484 81 13.7 Putative HC-toxin efflux carrier TOXA OS=Paenibacillus sp. TI45-13ar OX=1886670 GN=PTI45_04288 PE=3 SV=1
A0A1C5NQH3 444 81 13.7 Toxin A OS=uncultured Clostridium sp. OX=59620 GN=toxA_3 PE=4 SV=1
A0A1B7Y988 677 93 13.7 Cercosporin toxin biosynthesis protein OS=Colletotrichum higginsianum (strain IMI 349063) OX=759273 GN=CH
A0A194W677 568 81 13.7 Putative HC-toxin efflux carrier TOXA OS=Valsa mali OX=105487 GN=VM1G_07455 PE=4 SV=1
A0A194W5B1 588 81 13.7 Putative HC-toxin efflux carrier TOXA OS=Valsa mali OX=105487 GN=VM1G_07022 PE=4 SV=1
A0A174WZX2 555 81 13.7 Pertussis toxin liberation protein H OS=Fusicatenibacter sp. 2789STDY5834925 OX=1806509 GN=ptlH_3 PE=4 SV
A0A167PTC3 553 81 13.7 Putative HC-toxin efflux carrier TOXA OS=Penicillium chrysogenum OX=5076 GN=EN45_108990 PE=4 SV=1
A0A161XJV5 522 81 13.7 Zeta toxin OS=Pseudovibrio sp. WM33 OX=1735585 GN=PsWM33_01885 PE=4 SV=1
A0A160ULM8 464 81 13.7 Multidrug and toxin extrusion MATE family efflux pump YdhENorM homolog CDS OS=Bradyrhizobium sp. G22 OX=1
A0A158B7M0 457 81 13.7 Toxin-related secretion protein OS=Caballeronia hypogeia OX=1777140 GN=AWB79_03283 PE=4 SV=1
A0A146N1F7 481 81 13.7 Anthrax toxin receptor 1 OS=Fundulus heteroclitus OX=8078 PE=4 SV=1
A0A0U3TH36 438 81 13.7 Toxin HipA OS=Pantoea vagans OX=470934 GN=LK04_06180 PE=4 SV=1
A0A0U1M9S5 454 81 13.7 Putative HC-toxin efflux carrier TOXA OS=Talaromyces islandicus OX=28573 GN=PISL3812_09365 PE=4 SV=1
A0A0P9SF19 552 81 13.7 Putative cysteine peptidase toxin OS=Pseudomonas savastanoi pv. glycinea OX=318 GN=ALO37_02774 PE=4 SV=1
A0A0N1HIE5 564 81 13.7 Putative HC-toxin efflux carrier TOXA OS=Phialophora attae OX=1664694 GN=AB675_319 PE=4 SV=1
A0A0M2LEK5 422 81 13.7 Toxin HipA OS=Comamonas testosteroni OX=285 GN=XA67_01055 PE=4 SV=1
A0A0L6JAL4 455 81 13.7 Rtx toxin hemolysin-type calcium-binding protein OS=Methylobacterium sp. ARG-1 OX=1692501 GN=AKJ13_09980
A0A0K6HFJ1 773 106 13.7 Ca2+-binding protein, RTX toxin-related OS=Chelatococcus sambhunathii OX=363953 GN=Ga0061061_10983 PE=4
A0A0H4KVQ5 663 91 13.7 RTX toxin OS=Marinovum algicola DG 898 OX=988812 GN=MALG_01855 PE=4 SV=1
A0A0B1REV0 438 81 13.7 Toxin HipA OS=Pantoea rodasii OX=1076549 GN=QU24_01955 PE=4 SV=1
A0A068YKG4 619 85 13.7 Multidrug and toxin extrusion protein OS=Echinococcus multilocularis OX=6211 GN=EmuJ_001037000 PE=3 SV=1
U2F016 748 102 13.6 RTX toxin OS=Campylobacter concisus UNSW1 OX=1242967 GN=UNSW1_396 PE=4 SV=1
Q2S2P8 777 106 13.6 Probable toxin secretion ABC transporter ATP-binding protein, putative OS=Salinibacter ruber (strain DSM 138
H0TF61 649 88 13.6 Putative secretion protein (HlyD family) toxin/protease secretion system OS=Bradyrhizobium sp. STM 3843 OX=55
E3BLX2 639 87 13.6 RTX toxin RtxA-like protein OS=Vibrio caribbeanicus ATCC BAA-2122 OX=796620 GN=VIBC2010_01453 PE=4 SV=1
B6JI16 738 100 13.6 ABC transporter, toxin secretion OS=Oligotropha carboxidovorans (strain ATCC 49405 / DSM 1227 / KCTC 32145 /
A5EJN9 661 90 13.6 Putative toxin/protease secretion system OS=Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182) OX=288000 GN=BB
A0A2G5I8L5 627 85 13.6 Putative HC-toxin efflux carrier TOXA OS=Cercospora beticola OX=122368 GN=CB0940_00264 PE=4 SV=1
A0A1G5RCY2 785 107 13.6 Ca2+-binding protein, RTX toxin-related OS=Epibacterium ulvae OX=1156985 GN=SAMN04488118_11345 PE=4 SV=1
A0A1G5DNT0 726 99 13.6 Ca2+-binding protein, RTX toxin-related OS=Microvirga guangxiensis OX=549386 GN=SAMN02927923_00725 PE=4 S
A0A1C1CDE2 610 83 13.6 Putative HC-toxin efflux carrier TOXA OS=Cladophialophora carrionii OX=86049 GN=TOXA PE=4 SV=1
A0A1B2F8I5 714 97 13.6 Toxin RTX-I translocation ATP-binding protein OS=Pseudomonas putida OX=303 GN=apxIB_2 PE=4 SV=1
A0A199YQ35 656 89 13.6 RTX toxin OS=Lactococcus lactis RTB018 OX=1408188 GN=V425_12160 PE=4 SV=1
A0A034T9H9 634 86 13.6 RTX toxin and related Ca2+-binding protein OS=Vibrio sp. JCM 18904 OX=1298599 GN=JCM18904_1254 PE=4 SV=1
W9V5L6 728 98 13.5 RTX-I toxin determinant B OS=Nitrincola nitratireducens OX=1229521 GN=apxIB_1 PE=4 SV=1
U3JTL0 594 80 13.5 Multidrug and toxin extrusion protein OS=Ficedula albicollis OX=59894 GN=SLC47A2 PE=3 SV=1
S9RKQ0 446 80 13.5 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM,-like protein OS=Salipiger mucosus DSM 1609
Q9BZG8 443 80 13.5 2-(3-amino-3-carboxypropyl)histidine synthase subunit 1 OS=Homo sapiens OX=9606 GN=DPH1 PE=1 SV=2
Q75VA2 683 92 13.5 Putative mosquitocidal toxin OS=Bacillus thuringiensis subsp. entomocidus OX=1436 GN=cry30Aa like PE=3 SV=1
Q757B6 582 80 13.5 2-(3-amino-3-carboxypropyl)histidine synthase subunit 2 OS=Ashbya gossypii (strain ATCC 10895 / CBS 109.51 /
P45779 674 91 13.5 Type II secretion system protein D OS=Vibrio cholerae serotype O1 (strain ATCC 39315 / El Tor Inaba N16961) O
P0CN21 515 80 13.5 2-(3-amino-3-carboxypropyl)histidine synthase subunit 2 OS=Cryptococcus neoformans var. neoformans serotype D
P0CN20 515 80 13.5 2-(3-amino-3-carboxypropyl)histidine synthase subunit 2 OS=Cryptococcus neoformans var. neoformans serotype D
N1RUR3 549 80 13.5 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 4) OX=1229665 GN=FOC4
N0B285 440 80 13.5 Cholera toxin secretion EpsM protein OS=Hyphomicrobium denitrificans 1NES1 OX=670307 GN=HYPDE_29623 PE=4 SV=1
G4VQI1 734 99 13.5 Multidrug and toxin extrusion protein OS=Schistosoma mansoni OX=6183 GN=Smp_151290 PE=3 SV=1
C2HYF4 741 100 13.5 Cytolysin and hemolysin HlyA Pore-forming toxin OS=Vibrio albensis VL426 OX=593585 GN=VCA_000274 PE=4 SV=1
C1FUH4 750 101 13.5 Toxin complex component ORF-X2 OS=Clostridium botulinum (strain Kyoto / Type A2) OX=536232 GN=CLM_0892 PE=1
B3PH09 779 105 13.5 Putative toxin transporter OS=Cellvibrio japonicus (strain Ueda107) OX=498211 GN=CJA_3604 PE=4 SV=1
A6D2N0 540 80 13.5 Toxin secretion ABC transporter protein, HlyB family OS=Vibrio shilonii AK1 OX=391591 GN=VSAK1_25590 PE=4 SV=
A0A2I0LYI0 477 80 13.5 Multidrug and toxin extrusion protein OS=Columba livia OX=8932 GN=A306_00008031 PE=3 SV=1
A0A2G7DZ48 724 98 13.5 SUKH-4 immunity protein of toxin-antitoxin system OS=Streptomyces sp. 1 OX=492776 GN=CLW01_6727 PE=4 SV=1
A0A2G6YLJ3 739 100 13.5 SUKH-4 immunity protein of toxin-antitoxin system OS=Streptomyces sp. 76 OX=2035220 GN=CLW13_1428 PE=4 S

A0A2E6BDT9 434 80 13.5 Toxin HipA OS=Gammaproteobacteria bacterium OX=1913989 GN=CMQ46_03280 PE=4 SV=1
A0A2E1N9S7 401 80 13.5 Addiction module toxin YoeB OS=Halieaceae bacterium OX=2026743 GN=CME54_03710 PE=4 SV=1
A0A2D5T4E1 414 80 13.5 Toxin OS=Pseudoalteromonas sp. OX=53249 GN=CMK64_12580 PE=4 SV=1
A0A2C1TVY2 515 80 13.5 Mosquitocidal toxin protein OS=Bacillus cereus OX=1396 GN=CN271_31410 PE=4 SV=1
A0A2B5XG89 515 80 13.5 Mosquitocidal toxin protein OS=Bacillus wiedmannii OX=1890302 GN=CN611_28370 PE=4 SV=1
A0A2A7GGH4 492 80 13.5 Toxin OS=Bacillus thuringiensis OX=1428 GN=CON71_23755 PE=4 SV=1
A0A285EY80 555 80 13.5 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Actinoplanes atraurantiacus OX
A0A269YXP6 494 80 13.5 RTX toxin OS=Lactococcus lactis OX=1358 GN=B8W88_00110 PE=4 SV=1
A0A267EVH9 571 80 13.5 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig020709g1 PE=3 SV=1
A0A267EJH0 571 80 13.5 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig031687g1 PE=3 SV=1
A0A238VF34 487 80 13.5 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Hymenobacter mucosus OX=1411120 GN=SAMN06
A0A226WVS9 470 80 13.5 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Caballeronia sordidicola OX=196367 G
A0A218UU66 590 80 13.5 Multidrug and toxin extrusion protein OS=Lonchura striata domestica OX=299123 GN=SLC47A2 PE=3 SV=1
A0A1W4YCC3 631 85 13.5 Multidrug and toxin extrusion protein OS=Scleropages formosus OX=113540 GN=LOC108921227 PE=3 SV=1
A0A1Q8RA33 541 80 13.5 Putative HC-toxin efflux carrier TOXA 18 OS=Colletotrichum chlorophyti OX=708187 GN=CCHL11_07341 PE=4 SV=
A0A1N7M3Q7 687 93 13.5 Ca2+-binding protein, RTX toxin-related OS=Rhodobacter vinaykumarii OX=407234 GN=SAMN05421795_105150 PE=4
A0A1N6UVB0 621 84 13.5 Leukocidin/Hemolysin toxin family protein OS=Aeromonas sp. RU34C OX=1907417 GN=SAMN05880569_103389 PE=4 S
A0A1N6UHC0 621 84 13.5 Leukocidin/Hemolysin toxin family protein OS=Aeromonas hydrophila OX=644 GN=SAMN05878295_103389 PE=4 SV=1
A0A1N6QH73 637 86 13.5 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Luteimonas tolerans OX=1604334
A0A1K2ILP1 535 80 13.5 Metallopeptidase toxin 3 OS=Chryseobacterium limigenitum OX=1612149 GN=SAMN05216324_10542 PE=4 SV=1
A0A1I7N2Q1 474 80 13.5 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Hyphomicrobium facile OX=51670
A0A1I4VQP8 512 80 13.5 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Chryseobacterium oleae OX=491207 GN=SAMN0
A0A1I4J235 563 80 13.5 Ca2+-binding protein, RTX toxin-related OS=Thalassobius aestuarii OX=254406 GN=SAMN04488042_101824 PE=4 S
A0A1I1WSN9 917 124 13.5 Ca2+-binding protein, RTX toxin-related OS=Sulfitobacter brevis OX=74348 GN=SAMN04488523_104127 PE=4 SV=
A0A1I1MK79 447 80 13.5 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Flexibacter flexilis DSM 6793 OX=927664 G
A0A1H9P517 650 88 13.5 Papain fold toxin 1, glutamine deamidase OS=Lechevalieria xinjiangensis OX=402600 GN=SAMN05216188_111150
A0A1H8B4Y8 589 80 13.5 Predicted ribonuclease, toxin component of the YeeF-YezG toxin-antitoxin module OS=Paenisporosarcina quis
A0A1H7VF92 520 80 13.5 Ca2+-binding protein, RTX toxin-related OS=Syntrophus gentianae OX=43775 GN=SAMN04489760_103206 PE=4 SV=1
A0A1H6YWM2 508 80 13.5 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Dyadobacter koreensis OX=40865
A0A1H6W611 808 109 13.5 Ca2+-binding protein, RTX toxin-related OS=Sphingobium sp. AP50 OX=1884369 GN=SAMN05518849_10357 PE=4 SV
A0A1H3QMZ8 675 91 13.5 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Pseudomonas sp. NFIX28 OX=1566235 GN=SAMN
A0A1G8S7C0 786 106 13.5 Ca2+-binding protein, RTX toxin-related OS=Lutimaribacter saemankumensis OX=490829 GN=SAMN05421850_11013
A0A1G6S3K0 617 83 13.5 Ca2+-binding protein, RTX toxin-related OS=Ruegeria marina OX=639004 GN=SAMN04488239_105160 PE=4 SV=1
A0A1F8AGY3 418 80 13.5 Toxin biosynthesis protein (Tri7) OS=Aspergillus bombycis OX=109264 GN=ABOM_000504 PE=4 SV=1
A0A1D8ACI3 413 80 13.5 Toxin HipA OS=Novosphingobium resinovorum OX=158500 GN=BES08_23910 PE=4 SV=1
A0A1C6AER2 429 80 13.5 Toxin A OS=uncultured Clostridium sp. OX=59620 GN=toxA_2 PE=4 SV=1
A0A1C5TTT0 443 80 13.5 Toxin B OS=uncultured Clostridium sp. OX=59620 GN=toxB_2 PE=4 SV=1
A0A1C5SUI5 424 80 13.5 Pertussis toxin liberation protein H OS=uncultured Faecalibacterium sp. OX=259315 GN=ptlH PE=4 SV=1
A0A1C5PT50 443 80 13.5 Toxin B OS=uncultured Clostridium sp. OX=59620 GN=toxB_2 PE=4 SV=1
A0A1C5MYM4 518 80 13.5 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_1 PE=4 SV=1
A0A1C5KGM0 538 80 13.5 Toxin B OS=uncultured Clostridium sp. OX=59620 GN=toxB PE=4 SV=1
A0A1C1WRD2 448 80 13.5 Cercosporin toxin biosynthesis protein OS=Diaporthe helianthi OX=158607 GN=DHEL01_11641 PE=4 SV=1
A0A1C1CA27 589 80 13.5 Putative HC-toxin efflux carrier TOXA OS=Cladophialophora carrionii OX=86049 GN=TOXA PE=4 SV=1
A0A1C1C7V7 542 80 13.5 Putative HC-toxin efflux carrier TOXA OS=Cladophialophora carrionii OX=86049 GN=TOXA PE=4 SV=1
A0A1B5D875 418 80 13.5 Insecticidal Crystal Toxin, P42 OS=Pseudomonas sp. 44 R 15 OX=1844105 PE=4 SV=1
A0A1B4N5V0 440 80 13.5 Toxin HipA OS=Burkholderia multivorans OX=87883 GN=WK22_29180 PE=4 SV=1
A0A1A5HUJ3 412 80 13.5 Toxin HipA OS=Rhizobium loti OX=381 GN=BAE39_20860 PE=4 SV=1
A0A194VT73 562 80 13.5 Putative HC-toxin efflux carrier TOXA OS=Valsa mali OX=105487 GN=VM1G_03346 PE=4 SV=1
A0A194V6A7 568 80 13.5 Putative HC-toxin efflux carrier TOXA OS=Valsa mali var. pyri OX=694573 GN=VP1G_06765 PE=4 SV=1
A0A194V2V3 516 80 13.5 Putative HC-toxin efflux carrier TOXA OS=Valsa mali var. pyri OX=694573 GN=VP1G_11007 PE=4 SV=1
A0A175WOK0 588 80 13.5 Putative HC-toxin efflux carrier TOXA OS=Madurella mycetomatis OX=100816 GN=MMYC01_205596 PE=4 SV=1
A0A165VAU3 522 80 13.5 Zeta toxin OS=Pseudovibrio sp. Ad26 OX=989410 GN=PsAD26_01385 PE=4 SV=1
A0A158B8N3 450 80 13.5 Toxin HipA OS=Caballeronia catudaia OX=1777136 GN=AWB75_03116 PE=4 SV=1
A0A149PGO1 441 80 13.5 Toxin HipA OS=Paraburkholderia monticola OX=1399968 GN=CI15_26130 PE=4 SV=1
A0A135I3F7 589 80 13.5 RTX toxin OS=Enterovibrio coralii OX=294935 GN=ATN88_12070 PE=4 SV=1
A0A128EU77 588 80 13.5 RTX-I toxin determinant A from serotypes 1/9 OS=Grimontia celer OX=1796497 GN=apxIA PE=4 SV=1
A0A127MVH6 720 97 13.5 ABC transporter OS=Pseudomonas citronellolis OX=53408 GN=apxIB_4 PE=4 SV=1
A0A109LMF8 545 80 13.5 Dermonecrotic toxin OS=Pseudomonas fluorescens OX=294 GN=toxA_2 PE=4 SV=1
A0A109LCF4 815 110 13.5 Dermonecrotic toxin OS=Pseudomonas fluorescens OX=294 GN=toxA_4 PE=4 SV=1
A0A0N1DJG3 415 80 13.5 Phosphatidylinositol kinase OS=Rhizobium acidisoli OX=1538158 GN=AOG23_33430 PE=4 SV=1
A0A0N0JCN9 458 80 13.5 Toxin HipA OS=beta proteobacterium AAP65 OX=1523424 GN=IP80_10415 PE=4 SV=1
A0A0M8JIW9 465 80 13.5 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM, homolog OS=Sphingopyxis sp. C-1 OX=262
A0A0H6TQ70 561 80 13.5 RTX toxin transporter OS=Vibrio cholerae OX=666 GN=rtxB PE=4 SV=1
A0A0G9FJA5 443 80 13.5 Putative DNA-binding transcriptional regulator OS=Escherichia coli OX=562 GN=yjjJ PE=4 SV=1
A0A0G2ZS42 457 80 13.5 Multidrug and toxin extrusion OS=Archangium gephyra OX=48 GN=AA314_03180 PE=4 SV=1
A0A0F6W3W9 588 80 13.5 RTX toxin OS=Sandaracinus amylolyticus OX=927083 GN=DB32_003986 PE=4 SV=1
A0A0E9FUA5 461 80 13.5 Toxin B OS=Chlamydia trachomatis OX=813 GN=toxB PE=4 SV=1
A0A0C4YB24 450 80 13.5 Multidrug and toxin extrusion (MATE) family efflux pump, YdhE/NorM-like OS=Cupriavidus basilensis OX=6889

A0A0C1G3G3 434 80 13.5 RTX toxin OS=Ruegeria sp. ANG-R OX=1577903 GN=RA27_22010 PE=4 SV=1
A0A095RWM2 805 109 13.5 Cytolethal distending toxin A/C family protein OS=Burkholderia pseudomallei OX=28450 GN=DP49_2779 PE=4 S
A0A095DHL1 544 80 13.5 Antitoxin Phd YefM, type II toxin-antitoxin system family protein OS=Burkholderia cepacia OX=292 GN=DM43_
A0A093W0M5 547 80 13.5 Putative HC-toxin efflux carrier TOXA OS=Talaromyces marneffei PM1 OX=1077442 GN=GQ26_0032930 PE=4 SV=1
A0A090VFQ8 502 80 13.5 Putative RTX toxin hemolysin-type calcium-binding protein OS=Algibacter lectus OX=221126 GN=JCM19300_1959
A0A090T2C9 447 80 13.5 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM homolog OS=Vibrio maritimus OX=990268 G
A0A088Z8X6 802 108 13.5 Cytolethal distending toxin A/C family protein OS=Burkholderia pseudomallei OX=28450 GN=DP46_265 PE=4 SV
A0A061Q3K5 513 80 13.5 RTX toxin OS=Vibrio sp. JCM 19052 OX=1460647 GN=JCM19052_2254 PE=4 SV=1
X5V9I8 418 79 13.4 Toxin HipA OS=Mesorhizobium sp. LSHC422A00 OX=1287294 GN=X760_23860 PE=4 SV=1
W4HED1 486 79 13.4 RTX toxin OS=Roseivivax atlanticus OX=1317118 GN=ATO8_18974 PE=4 SV=1
U6PX46 544 79 13.4 Tyrosinase and Metridin ShK toxin domain containing protein OS=Haemonchus contortus OX=6289 GN=HCOI_00065800
U3JBR6 492 79 13.4 Anthrax toxin receptor 1 OS=Ficedula albicollis OX=59894 GN=ANTXR1 PE=4 SV=1
U1S392 572 79 13.4 Toxin secretion/phage lysis holin OS=Actinomyces johnsonii F0542 OX=1321818 GN=HMPREF1979_00797 PE=3 SV=1
R9UMC7 469 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Monascus pilosus OX=89488 GN=2383 PE=4 SV=1
Q9H0J9 701 94 13.4 Poly [ADP-ribose] polymerase 12 OS=Homo sapiens OX=9606 GN=PARP12 PE=1 SV=1
Q5NV03 605 81 13.4 Hypothetical Toxin corregulated pilus biosynthesis protein T OS=Cupriavidus metallidurans (strain ATCC 43123
Q4WN99 565 79 13.4 2-(3-amino-3-carboxypropyl)histidine synthase subunit 2 OS=Neosartorya fumigata (strain ATCC MYA-4609 / Af293
Q3RGQ1 444 79 13.4 Zonular occludens toxin OS=Xylella fastidiosa Dixon OX=155919 GN=XfasaDRAFT_2260 PE=4 SV=1
Q3ES64 501 79 13.4 MOSQUITOCIDAL TOXIN PROTEIN OS=Bacillus thuringiensis serovar israelensis ATCC 35646 OX=339854 GN=RBTH_04010
N4UYG3 504 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 1) OX=1229664 GN=FOC1
N1S2A6 529 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 4) OX=1229665 GN=FOC4
N1RYR6 565 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 4) OX=1229665 GN=FOC4
L7ISW9 659 88 13.4 Multidrug and toxin extrusion protein 1 OS=Magnaporthe oryzae (strain P131) OX=1143193 GN=OOW_P131scaffold013
L7I120 659 88 13.4 Multidrug and toxin extrusion protein 1 OS=Magnaporthe oryzae (strain Y34) OX=1143189 GN=OOU_Y34scaffold00629
K2QIF3 413 79 13.4 Toxin biosynthesis protein (Tri7), putative OS=Macrophomina phaseolina (strain MS6) OX=1126212 GN=MPH_13396 P
I3K943 499 79 13.4 Anthrax toxin receptor 2a OS=Oreochromis niloticus OX=8128 GN=antxr2 PE=4 SV=1
H3DFK3 484 79 13.4 Multidrug and toxin extrusion protein OS=Tetraodon nigroviridis OX=99883 PE=3 SV=1
H0TF60 583 79 13.4 Putative secretion ATP-binding protein (ABC-type transporter family) toxin/protease secretion system OS=Brady
G4NF92 659 88 13.4 Multidrug and toxin extrusion protein 1 OS=Magnaporthe oryzae (strain 70-15 / ATCC MYA-4617 / FGSC 8958) OX=2
F7J0A4 454 79 13.4 Iota toxin component Ia OS=Clostridium perfringens OX=1502 GN=iap PE=4 SV=1
F4UX72 457 79 13.4 Putative toxin-antitoxin system, toxin component OS=Escherichia coli TA280 OX=656444 GN=ECNG_01420 PE=4 SV=1
E2CST8 416 79 13.4 RTX-III toxin determinant A from serotype 2 OS=Roseibium sp. TrichSKD4 OX=744980 GN=TRICHSKD4_6243 PE=4 SV=1
C5UY10 487 79 13.4 Toxin complex component ORF-X3 OS=Clostridium botulinum E1 str. 'BoNT E Beluga' OX=536233 GN=CLO_2647 PE=4 SV
C4IHL5 487 79 13.4 Toxin complex component ORF-X3 OS=Clostridium butyricum E4 str. BoNT E BL5262 OX=632245 GN=CLP_2746 PE=4 SV=1
B8CUM8 552 79 13.4 Toxin secretion ABC transporter protein, HlyB family OS=Shewanella piezotolerans (strain WP3 / JCM 13877) OX=
B0Y6I6 534 79 13.4 Toxin biosynthesis cytochrome P450 monooxygenase, putative OS=Neosartorya fumigata (strain CEA10 / CBS 144.89
A5VLV0 369 79 13.4 Membrane protein putative toxin regulator-like protein OS=Lactobacillus reuteri (strain DSM 20016) OX=557436
A3JRR9 574 79 13.4 ABC protein toxin exporter, fused ATPase and inner membrane domain OS=Rhodobacteraceae bacterium HTCC2150 OX=
A0A2K4W2B4 504 79 13.4 Zeta toxin family protein OS=Pseudomonas cerasi OX=1583341 GN=PL963_P100038 PE=4 SV=1
A0A2J9F9P6 445 79 13.4 Type II toxin-antitoxin system HipA family toxin OS=Yersinia enterocolitica OX=630 GN=CEQ35_008115 PE=4 S
A0A2IOGW32 364 79 13.4 Toxin regulator PfoR OS=Psychrobacter sp. 4Bb OX=888436 GN=CXF60_09540 PE=4 SV=1
A0A2G7F6Y6 687 92 13.4 SUKH-4 immunity protein of toxin-antitoxin system OS=Streptomyces sp. 70 OX=1827606 GN=CLW09_07158 PE=4 S
A0A2G6XRH7 443 79 13.4 Nucleic acid/nucleotide deaminase of polymorphic system toxin OS=Streptomyces sp. 76 OX=2035220 GN=CLW13_
A0A2G6II66 435 79 13.4 RTX toxin OS=Rhodobacterales bacterium OX=1948890 GN=CSA72_07055 PE=4 SV=1
A0A2G5I2Y5 571 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Cercospora beticola OX=122368 GN=CB0940_02507 PE=4 SV=1
A0A2G4E4B7 504 79 13.4 Toxin OS=Pseudomonas sp. NZIPFR-PS5 OX=1230465 GN=AO263_17565 PE=4 SV=1
A0A2G2I6F3 459 79 13.4 Zonular occludens toxin OS=Colwellia sp. OX=56799 GN=COA59_15595 PE=4 SV=1
A0A2E4TQP6 433 79 13.4 Type II toxin-antitoxin system HipA family toxin OS=Confluentimicrobium sp. OX=1931204 GN=CL813_15680 PE=
A0A2D3RC24 434 79 13.4 Toxin HipA OS=Sphingorhabdus flavimaris OX=266812 GN=CHN51_06625 PE=4 SV=1
A0A2B3XY35 516 79 13.4 Mosquitocidal toxin protein OS=Bacillus anthracis OX=1392 GN=COK92_26115 PE=4 SV=1
A0A2A9FI53 581 79 13.4 Putative AbiEii toxin of type IV toxin-antitoxin system OS=Amycolatopsis sulphurea OX=76022 GN=ATK36_6327
A0A286JZR1 504 79 13.4 Zeta toxin OS=Pseudomonas syringae pv. actinidiae OX=103796 PE=4 SV=1
A0A271KXY6 441 79 13.4 Toxin HipA OS=Mesorhizobium mediterraneum OX=43617 GN=CIT25_20100 PE=4 SV=1
A0A267EBX3 633 85 13.4 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig032378g2 PE=3 SV=1
A0A1Y6JNR4 723 97 13.4 Multidrug/toxin efflux protein, ATP-binding and membrane protein OS=Pseudomonas viridiflava OX=33069 GN=C
A0A1X4N7D0 444 79 13.4 Toxin HipA OS=Thalassospira sp. MCCC 1A01428 OX=1470575 GN=THS27_15635 PE=4 SV=1
A0A1XOMGF2 446 79 13.4 Toxin HipA OS=Burkholderia sp. A27 OX=1755991 GN=B2G74_29145 PE=4 SV=1
A0A1V5TY67 584 79 13.4 Toxin RTX-I translocation ATP-binding protein OS=Firmicutes bacterium ADurb.Bin248 OX=1852886 GN=apxIB PE
A0A1VORKW9 738 99 13.4 Toxin RTX-I translocation ATP-binding protein OS=Roseovarius mucosus OX=215743 GN=apxIB PE=4 SV=1
A0A1VOB919 444 79 13.4 Toxin HipA OS=Pseudomonas sp. S-6-2 OX=1931241 GN=BVH74_17500 PE=4 SV=1
A0A1R3WCY1 598 80 13.4 Papain fold toxin 1, glutamine deamidase OS=Loktanella rosea OX=287098 GN=SAMN05421665_0246 PE=4 SV=1
A0A1R3UML7 643 86 13.4 RTX toxin transporter, ATP-binding protein OS=Nocardiopsis sp. JB363 OX=1434837 GN=BQ8420_15405 PE=4 SV=1
A0A1R2D5R0 444 79 13.4 Zonular occludens toxin OS=Xylella fastidiosa subsp. multiplex OX=644357 GN=XYFPCFBP8417_10195 PE=4 SV=1
A0A1Q8RZP8 596 80 13.4 Putative HC-toxin efflux carrier TOXA 22 OS=Colletotrichum chlorophyti OX=708187 GN=CCHL11_07579 PE=4 SV=
A0A1N7HPC5 406 79 13.4 Ca2+-binding protein, RTX toxin-related OS=Roseovarius nanhaiticus OX=573024 GN=SAMN05421666_3597 PE=4 SV
A0A1M6MJW3 514 79 13.4 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Halomonas sinaiensis OX=379597
A0A1LOD980 613 82 13.4 Toxin OS=Moritella viscosa OX=80854 GN=NVI5450_0249 PE=4 SV=1
A0A1K1M2P8 598 80 13.4 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Prevotellaceae bacterium HUN15

A0A1J5N226 732 98 13.4 Toxin RTX-I translocation ATP-binding protein OS=Desulfovibrio dechloracetivorans OX=117209 GN=apxIB PE=4
A0A1I9YUU3 441 79 13.4 Toxin HipA OS=Paraburkholderia sprentiae WSM5005 OX=754502 GN=BJG93_31660 PE=4 SV=1
A0A1I8HDZ6 633 85 13.4 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 PE=3 SV=1
A0A1I7A8G8 707 95 13.4 Ca2+-binding protein, RTX toxin-related OS=Mesorhizobium sp. YR577 OX=1884373 GN=SAMN05518861_10441 PE=4
A0A1I6YG41 522 79 13.4 Ca2+-binding protein, RTX toxin-related OS=Mesorhizobium sp. YR577 OX=1884373 GN=SAMN05518861_101556 PE=4
A0A1I3YAJ4 433 79 13.4 Ca2+-binding protein, RTX toxin-related OS=Methylobacterium salsuginis OX=414703 GN=SAMN04488125_10192 PE
A0A1H8FHW5 404 79 13.4 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Prevotella sp. ne3005 OX=17618
A0A1H4ZHL0 460 79 13.4 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Pseudomonas proteolytica OX=21
A0A1H1WV39 542 79 13.4 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Pseudomonas sp. bs2935 OX=1761
A0A1D5QIQ9 449 79 13.4 Multidrug and toxin extrusion protein OS=Macaca mulatta OX=9544 GN=SLC47A1 PE=3 SV=1
A0A1C6M6Z3 581 79 13.4 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Streptomyces sp. LamerLS-316 O
A0A1C6L0D2 506 79 13.4 Toxin A OS=uncultured Clostridium sp. OX=59620 GN=toxA_3 PE=4 SV=1
A0A1C6DG68 506 79 13.4 Toxin A OS=uncultured Ruminococcus sp. OX=165186 GN=toxA_2 PE=4 SV=1
A0A1C6D3T2 506 79 13.4 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_1 PE=4 SV=1
A0A1C5LFT3 501 79 13.4 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_1 PE=4 SV=1
A0A1C5KNW4 828 111 13.4 Toxin A OS=uncultured Clostridium sp. OX=59620 GN=toxA_8 PE=4 SV=1
A0A1C4CGN7 498 79 13.4 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Kosakonia oryzendophytica OX=1
A0A1B2D4G5 435 79 13.4 Toxin HipA OS=Comamonas aquatica OX=225991 GN=MA05_11430 PE=4 SV=1
A0A1B1V690 604 81 13.4 Cytolysin and hemolysin HlyA Pore-forming toxin OS=Vibrio coralliilyticus OX=190893 GN=BA953_00935 PE=4 S
A0A1B1Q459 414 79 13.4 Pertussis toxin liberation protein G OS=Prosthecochloris sp. CIB 2401 OX=1868325 GN=ptlG PE=4 SV=1
A0A1A7V131 510 79 13.4 Toxin coregulated pilus biosynthesis protein T OS=Vibrio mediterranei OX=689 GN=tcpT PE=4 SV=1
A0A196QOL7 433 79 13.4 Toxin HipA OS=Sulfitobacter geojensis OX=1342299 GN=A8B74_08605 PE=4 SV=1
A0A175WHF4 579 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Madurella mycetomatis OX=100816 GN=MMYC01_200288 PE=4 SV=1
A0A174X9K4 506 79 13.4 Toxin A OS=[Ruminococcus] torques OX=33039 GN=toxA_3 PE=4 SV=1
A0A174KPZ8 568 79 13.4 Toxin A OS=Dorea longicatena OX=88431 GN=toxA_5 PE=4 SV=1
A0A161XWH5 499 79 13.4 Hc-toxin efflux carrier toxa OS=Colletotrichum incanum OX=1573173 GN=CI238_13203 PE=4 SV=1
A0A158KLC4 457 79 13.4 Toxin HipA OS=Caballeronia arvi OX=1777135 GN=AWB74_05933 PE=4 SV=1
A0A158BYS2 457 79 13.4 Toxin-related secretion protein OS=Caballeronia glebae OX=1777143 GN=AWB82_04764 PE=4 SV=1
A0A151MQ65 496 79 13.4 Anthrax toxin receptor 2 OS=Alligator mississippiensis OX=8496 GN=ANTXR2 PE=4 SV=1
A0A150XF36 425 79 13.4 Toxin HipA OS=Roseivirga spongicola OX=333140 GN=AWW68_00715 PE=4 SV=1
A0A146Y1N7 687 92 13.4 Multidrug and toxin extrusion protein OS=Fundulus heteroclitus OX=8078 PE=3 SV=1
A0A142LIF1 430 79 13.4 Toxin HipA OS=Betaproteobacteria bacterium UKL13-2 OX=1690485 GN=AEM42_05485 PE=4 SV=1
A0A142JIK1 448 79 13.4 Toxin HipA OS=Cupriavidus nantongensis OX=1796606 GN=A2G96_09270 PE=4 SV=1
A0A108UBF2 477 79 13.4 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Lysobacter capsici AZ78 OX=1444315 G
A0A0U1LYH4 585 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Talaromyces islandicus OX=28573 GN=PISL3812_05417 PE=4 SV=1
A0A0U1LXH7 548 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Talaromyces islandicus OX=28573 GN=PISL3812_05096 PE=4 SV=1
A0A0U1HFH8 710 95 13.4 Putative toxin transport protein OS=Yersinia enterocolitica OX=630 GN=apxIB_2 PE=4 SV=1
A0A0S7DZ11 642 86 13.4 Putative HC-toxin efflux carrier TOXA OS=Aspergillus lentulus OX=293939 GN=ALT_5401 PE=4 SV=1
A0A0R3LAW0 413 79 13.4 Toxin HipA OS=Bradyrhizobium jicamae OX=280332 GN=CQ12_18670 PE=4 SV=1
A0A0R2JTS8 347 79 13.4 Membrane protein, putative toxin regulator OS=Lactobacillus lindneri DSM 20690 = JCM 11027 OX=1122148 GN=
A0A0Q7KYE3 445 79 13.4 Toxin HipA OS=Variovorax sp. Root473 OX=1736541 GN=ASD34_16255 PE=4 SV=1
A0A0P9TKE5 423 79 13.4 Zonular occludens toxin OS=Pseudomonas savastanoi pv. glycinea OX=318 GN=ALO37_02716 PE=4 SV=1
A0A0P9MDT2 423 79 13.4 PbsX-like transcriptional regulator OS=Pseudomonas syringae pv. cerasicola OX=264451 GN=ALO50_200077 PE=4
A0A0P7AS97 544 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Neonectria ditissima OX=78410 GN=AK830_g9167 PE=4 SV=1
A0A0P1GY11 441 79 13.4 Pertussis toxin liberation protein G OS=Tropicibacter naphthalenivorans OX=441103 GN=ptlG PE=4 SV=1
A0A0PORCX4 482 79 13.4 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Paraburkholderia caribensis MBA4 OX=
A0A0N1HD37 564 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Phialophora attae OX=1664694 GN=AB675_2170 PE=4 SV=1
A0A0M9TXY5 486 79 13.4 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM, homolog OS=Sphingopyxis sp. C-1 OX=262
A0A0H1RNI1 657 88 13.4 RTX toxin OS=Lactococcus lactis subsp. lactis OX=1360 GN=VN91_0923 PE=4 SV=1
A0A0F7KUX1 511 79 13.4 Zeta toxin OS=Altererythrobacter atlanticus OX=1267766 GN=WYH_01936 PE=4 SV=1
A0A0D6DUG1 525 79 13.4 Putative LXG domain-containing toxin OS=Lactococcus piscium MKFS47 OX=297352 GN=LACPI_0357 PE=4 SV=1
A0A0C2SN97 487 79 13.4 OrfX3 protein OS=Clostridium botulinum OX=1491 GN=ADT22_04995 PE=4 SV=1
A0A0B8Q8Q3 487 79 13.4 Multidrug and toxin extrusion family efflux pump ydhE/norM OS=Vibrio ishigakensis OX=1481914 GN=JCM19241_
A0A0B8P3T8 426 79 13.4 Multidrug and toxin extrusion family efflux pump ydhE/norM OS=Vibrio ishigakensis OX=1481914 GN=JCM19232_
A0A0B7J0Q2 745 100 13.4 RTX-I toxin determinant B OS=Candidatus Methylopumilus turicensis OX=1581680 GN=apxIB PE=4 SV=1
A0A096A930 419 79 13.4 Toxin HipA OS=Prevotella melaninogenica DNF00666 OX=1401073 GN=HMPREF0661_11870 PE=4 SV=1
A0A093V816 546 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Talaromyces marneffei PM1 OX=1077442 GN=GQ26_0120500 PE=4 SV=1
A0A093V079 578 79 13.4 Putative HC-toxin efflux carrier TOXA OS=Talaromyces marneffei PM1 OX=1077442 GN=GQ26_0231690 PE=4 SV=1
A0A090R555 448 79 13.4 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM homolog OS=Vibrio sp. C7 OX=1001886 GN=
A0A085WEI8 467 79 13.4 Multidrug and toxin extrusion (MATE) family efflux pump protein OS=Hyalangium minutum OX=394096 GN=DB31_1
A0A074XVY8 428 79 13.4 Putative toxin biosynthesis protein OS=Aureobasidium pullulans EXF-150 OX=1043002 GN=M438DRAFT_393631 PE=
Q1I6U8 714 95 13.3 Putative multidrug/toxin efflux protein, ATP binding and membrane protein OS=Pseudomonas entomophila (strain
K7G9A8 600 80 13.3 Multidrug and toxin extrusion protein OS=Pelodiscus sinensis OX=13735 GN=SLC47A2 PE=3 SV=1
J9GM92 656 87 13.3 Zonular occludens toxin OS=gut metagenome OX=749906 GN=EVA_03272 PE=4 SV=1
A0A1X7VPD1 626 83 13.3 Multidrug and toxin extrusion protein OS=Amphimedon queenslandica OX=400682 PE=3 SV=1
A0A1T4ZJB3 817 109 13.3 Hemolysin-type calcium-binding toxin (Secreted) OS=Planktothrix sp. PCC 11201 OX=1729650 GN=PL11201_6300
A0A1S3NSZ6 607 81 13.3 Multidrug and toxin extrusion protein OS=Salmo salar OX=8030 GN=LOC106581174 PE=3 SV=1
A0A1Q5TU82 698 93 13.3 RTX toxin OS=Xenorhabdus thuongxuanensis OX=1873484 GN=Xentx_02909 PE=4 SV=1

A0A1L9P093 707 94 13.3 Toxin RTX-I translocation ATP-binding protein OS=Planktotalea frisia OX=696762 GN=apxIB PE=4 SV=1
A0A1K1S5P1 690 92 13.3 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Pseudomonas sp. NFACC16-2 OX=1554560 GN=S
A0A1H6H0C7 791 105 13.3 Ca2+-binding protein, RTX toxin-related OS=Selenomonas ruminantium OX=971 GN=SAMN05216583_11048 PE=4 SV=
A0A1H0JTI2 624 83 13.3 Ca2+-binding protein, RTX toxin-related OS=Albidiferax sp. OV413 OX=1855285 GN=SAMN05216303_1011486 PE=4
A0A1G7MAV6 622 83 13.3 Ca2+-binding protein, RTX toxin-related OS=Rhodobacter capsulatus OX=1061 GN=SAMN04244550_02488 PE=4 SV=1
A0A1C6H682 707 94 13.3 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_2 PE=4 SV=1
A0A0S2GBR1 678 90 13.3 Insecticide toxin TcdB middle/N-terminal region family protein OS=Lysobacter gummosus OX=262324 GN=LG3211
A0A0K0Y460 754 100 13.3 Toxin RTX-I translocation ATP-binding protein OS=Octadecabacter temperatus OX=1458307 GN=apxIB PE=4 SV=1
A0A0F6LGC3 800 106 13.3 Cytolethal distending toxin A/C family protein OS=Burkholderia pseudomallei MSHR4000 OX=1435370 GN=X980_
A0A0E3DRZ0 800 106 13.3 Putative Rhs family protein OS=Escherichia coli OX=562 GN=MS6198_A150 PE=4 SV=1
A0A077PIY4 639 85 13.3 C component of insecticidal toxin complex OS=Xenorhabdus bovienii str. kraussei Quebec OX=1398203 GN=XBKQ
A0A034TMS2 747 99 13.3 RTX toxin OS=Vibrio sp. JCM 18904 OX=1298599 GN=JCM18904_5104 PE=4 SV=1
A0A021XII4 624 83 13.3 RTX toxin OS=Shinella sp. DD12 OX=1410620 GN=SHLA_61c000450 PE=4 SV=1
W7WYL0 576 78 13.2 Pertussis toxin liberation protein C OS=Hydrogenophaga sp. T4 OX=1437444 GN=ptlC PE=4 SV=1
W6U184 420 78 13.2 Multidrug and toxin extrusion protein OS=Echinococcus granulosus OX=6210 GN=EGR_10894 PE=4 SV=1
W5L2G5 483 78 13.2 Anthrax toxin receptor 2b OS=Astyanax mexicanus OX=7994 PE=4 SV=1
W2ECP1 416 78 13.2 Toxin-like protein OS=Paenibacillus larvae subsp. larvae DSM 25719 OX=697286 GN=ERIC1_1c29940 PE=4 SV=1
W1J328 713 94 13.2 Toxin RTX-I translocation ATP-binding protein OS=Xenorhabdus szentirmaii DSM 16338 OX=1427518 GN=apxIB PE=4 S
V9W001 416 78 13.2 Toxin-like protein OS=Paenibacillus larvae subsp. larvae DSM 25430 OX=697284 GN=ERIC2_c03770 PE=4 SV=1
V7HF01 418 78 13.2 Toxin HipA OS=Mesorhizobium sp. L103C120A0 OX=1287086 GN=X728_26525 PE=4 SV=1
V5F763 615 81 13.2 Toxin OS=Photobacterium leiognathi lrivu.4.1 OX=1248232 GN=PLEI_3644 PE=4 SV=1
T0PBV8 801 106 13.2 Insecticidal toxin complex protein TccC2 OS=Photorhabdus temperata subsp. temperata M1021 OX=1221520 GN=B738
S5FPJ4 445 78 13.2 HC-toxin bZIP transcription factor OS=Alternaria jesenskae OX=378183 GN=TOXE PE=2 SV=1
R7V497 591 78 13.2 Multidrug and toxin extrusion protein OS=Capitella teleta OX=283909 GN=CAPTEDRAFT_217910 PE=3 SV=1
R6YM58 472 78 13.2 Toxin-antitoxin system toxin component Fic family OS=Alistipes sp. CAG:435 OX=1262695 GN=BN655_01761 PE=4 SV=
Q96FL8-3 586 78 13.2 Isoform 3 of Multidrug and toxin extrusion protein 1 OS=Homo sapiens OX=9606 GN=SLC47A1
Q46220 454 78 13.2 Iota toxin component Ia OS=Clostridium perfringens OX=1502 PE=1 SV=1
Q3SYT1 438 78 13.2 2-(3-amino-3-carboxypropyl)histidine synthase subunit 1 OS=Bos taurus OX=9913 GN=DPH1 PE=2 SV=1
P0CN19 529 78 13.2 2-(3-amino-3-carboxypropyl)histidine synthase subunit 1 OS=Cryptococcus neoformans var. neoformans serotype D
P0CN18 529 78 13.2 2-(3-amino-3-carboxypropyl)histidine synthase subunit 1 OS=Cryptococcus neoformans var. neoformans serotype D
N1RVD4 582 78 13.2 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 4) OX=1229665 GN=FOC4
L8JDE7 535 78 13.2 RTX toxin transporter OS=Photobacterium marinum OX=1056511 GN=C942_04570 PE=4 SV=1
K6XJ16 558 78 13.2 Toxin secretion ABC transporter ATP-binding protein OS=Paraglaciecola agarilytica NO2 OX=1125747 GN=GAGA_3238
H3BGB0 484 78 13.2 Anthrax toxin receptor 1 OS=Latimeria chalumnae OX=7897 GN=ANTXR1 PE=4 SV=1
H2RBV0 586 78 13.2 Multidrug and toxin extrusion protein OS=Pan troglodytes OX=9598 GN=SLC47A1 PE=3 SV=1
F4AKV8 558 78 13.2 Toxin secretion ABC transporter ATP-binding protein OS=Glaciecola sp. (strain 4H-3-7+YE-5) OX=983545 GN=Glaag
F2PW03 420 78 13.2 Cercosporin toxin biosynthesis protein OS=Trichophyton equinum (strain ATCC MYA-4606 / CBS 127.97) OX=559882
E9ESD7 496 78 13.2 Multidrug and toxin extrusion protein 1 OS=Metarhizium robertsii (strain ARSEF 23 / ATCC MYA-3075) OX=655844
D8JJL9 707 93 13.2 RTX toxin OS=Acinetobacter oleivorans (strain JCM 16667 / KCTC 23045 / DR1) OX=436717 GN=AOLE_09055 PE=4 SV=1
D0HE00 743 98 13.2 Cytolysin and hemolysin HlyA Pore-forming toxin OS=Vibrio mimicus VM223 OX=675820 GN=VMA_001263 PE=4 SV=1
C8VQ97 620 82 13.2 MFS toxin efflux pump (AflT), putative (AFU_orthologue AFUA_1G12620) OS=Emericella nidulans (strain FGSC A4 /
B9JJE1 479 78 13.2 Rhizobiocin/RTX toxin and hemolysin-type calcium binding protein OS=Agrobacterium radiobacter (strain K84 / A
B1L2F9 490 78 13.2 Toxin complex component ORF-X3 OS=Clostridium botulinum (strain Loch Maree / Type A3) OX=498214 GN=CLK_A0070
B1BSY1 454 78 13.2 Iota toxin component Ia OS=Clostridium perfringens E str. JGS1987 OX=451755 GN=AC3_A0576 PE=4 SV=1
A1CWW7 500 78 13.2 Toxin biosynthesis cytochrome P450 monooxygenase, putative OS=Neosartorya fischeri (strain ATCC 1020 / DSM 37
A0A2K5W799 442 78 13.2 Multidrug and toxin extrusion protein 1 OS=Macaca fascicularis OX=9541 PE=4 SV=1
A0A2I3GME6 568 78 13.2 Multidrug and toxin extrusion protein OS=Nomascus leucogenys OX=61853 GN=SLC47A2 PE=3 SV=1
A0A2I0GPB8 364 78 13.2 Toxin regulator PfoR OS=Psychrobacter sp. 4Dc OX=888437 GN=CXF61_06510 PE=4 SV=1
A0A2H3F5G5 667 88 13.2 Multidrug and toxin extrusion protein OS=Diplocarpon rosae OX=946125 GN=BUE80_DR011826 PE=4 SV=1
A0A2G5M4W2 446 78 13.2 Toxin HipA OS=Pseudomonas sp. 2588-5 OX=1712676 GN=AOA57_18160 PE=4 SV=1
A0A2D9YIJ0 453 78 13.2 RTX toxin OS=Maritimibacter sp. OX=2003363 GN=CMH12_01235 PE=4 SV=1
A0A2D7VGE4 443 78 13.2 Type II toxin-antitoxin system HipA family toxin OS=Acinetobacter sp. OX=472 GN=CL490_15185 PE=4 SV=1
A0A2D3IKA6 418 78 13.2 Toxin HipA OS=Phyllobacterium sp. Tri-48 OX=1867719 GN=BLM14_19635 PE=4 SV=1
A0A291E6F5 425 78 13.2 Type II toxin-antitoxin system HipA family toxin OS=Cedecea neteri OX=158822 GN=CO704_25900 PE=4 SV=1
A0A249NT10 466 78 13.2 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Sinorhizobium sp. CCBAU 05631 OX=794
A0A242MQ24 474 78 13.2 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Caballeronia sordidicola OX=196367 G
A0A210PYV3 631 83 13.2 Multidrug and toxin extrusion protein OS=Mizuhopecten yessoensis OX=6573 GN=KP79_PYT02145 PE=3 SV=1
A0A1Y5SHY9 748 99 13.2 Toxin RTX-I translocation ATP-binding protein OS=Roseisalinus antarcticus OX=254357 GN=apxIB PE=4 SV=1
A0A1Y5RJQ7 744 98 13.2 Toxin RTX-I translocation ATP-binding protein OS=Pseudooctadecabacter jejudonensis OX=1391910 GN=apxIB_1
A0A1U9Q5Z1 446 78 13.2 Toxin HipA OS=Pseudomonas azotoformans OX=47878 GN=B1R45_29865 PE=4 SV=1
A0A1U0JK20 349 78 13.2 Predicted membrane protein, putative toxin regulator OS=Mycobacterium abscessus subsp. abscessus OX=11856
A0A1S9GFH4 680 90 13.2 RTX toxin OS=Rhizobium leguminosarum bv. viciae USDA 2370 OX=754774 GN=BS629_31445 PE=4 SV=1
A0A1S2JS32 422 78 13.2 Toxin HipA OS=Pseudomonas putida OX=303 GN=BIW19_08600 PE=4 SV=1
A0A1R3EAF7 474 78 13.2 Zonular occludens toxin OS=Vibrio splendidus OX=29497 GN=BH581_23375 PE=4 SV=1
A0A1Q3E824 427 78 13.2 Cercosporin toxin biosynthesis protein OS=Lentinula edodes OX=5353 GN=LENED_004926 PE=4 SV=1
A0A1N7JQB0 513 78 13.2 Restriction endonuclease fold toxin 5 OS=Chryseobacterium shigense OX=297244 GN=SAMN05421639_10797 PE=4 S
A0A1J5L5L5 441 78 13.2 RTX toxin OS=Alphaproteobacteria bacterium MedPE-SWcel OX=1860092 GN=BM562_11675 PE=4 SV=1
A0A1I9YLP6 446 78 13.2 Toxin HipA OS=Paraburkholderia sprentiae WSM5005 OX=754502 GN=BJG93_16965 PE=4 SV=1

A0A1I7L4T9 404 78 13.2 Zona occludens toxin OS=Polaromonas sp. YR568 OX=1855301 GN=SAMN05216350_12110 PE=4 SV=1
A0A1I6IZK5 560 78 13.2 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Microbacterium sp. cl127 OX=17
A0A1I5GUU9 460 78 13.2 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Pseudomonas sp. NFACC24-1 OX=1
A0A1I3NNW4 727 96 13.2 Ca2+-binding protein, RTX toxin-related OS=Celeribacter halophilus OX=576117 GN=SAMN04488138_1027 PE=4 SV
A0A1I1BDM9 652 86 13.2 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Janthinobacterium sp. 344 OX=1
A0A1H7UA60 403 78 13.2 Zona occludens toxin OS=Roseateles sp. YR242 OX=1855305 GN=SAMN05216359_1296 PE=4 SV=1
A0A1H5R216 495 78 13.2 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Amycolatopsis pretoriensis OX=
A0A1H5J9I1 497 78 13.2 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Streptomyces sp. 2131.1 OX=185
A0A1H3IMG4 429 78 13.2 Papain fold toxin 1, glutamine deamidase OS=Saccharopolyspora shandongensis OX=418495 GN=SAMN05216215_102
A0A1H3H9C2 457 78 13.2 Toxin 24 OS=Pseudomonas syringae OX=317 GN=SAMN05444506_101448 PE=4 SV=1
A0A1H3DE47 474 78 13.2 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Hymenobacter psychrophilus OX=651662 GN=S
A0A1H2YMA3 486 78 13.2 Nucleotidyl transferase AbiEii toxin, Type IV TA system OS=Amycolatopsis xylanica OX=589385 GN=SAMN054215
A0A1HOWD34 758 100 13.2 Ca2+-binding protein, RTX toxin-related OS=Phyllobacterium sp. YR620 OX=1881066 GN=SAMN05428967_4110 PE=
A0A1G8I386 537 78 13.2 Toxin CptA OS=Arthrobacter cupressi OX=1045773 GN=SAMN05216555_101114 PE=3 SV=1
A0A1G7BQB4 484 78 13.2 Ca2+-binding protein, RTX toxin-related OS=Belnapia rosea OX=938405 GN=SAMN04487779_102626 PE=4 SV=1
A0A1G6F374 940 124 13.2 Insecticidal toxin complex protein TccC OS=Pseudomonas putida OX=303 GN=SAMN03097715_05463 PE=4 SV=1
A0A1G5MKJ1 940 124 13.2 Insecticidal toxin complex protein TccC OS=Pseudomonas sp. NFIX46 OX=1566234 GN=SAMN03159313_1949 PE=4 S
A0A1G3FG71 415 78 13.2 RTX toxin OS=Rhodobacteraceae bacterium GWE1_64_9 OX=1802012 GN=A2092_18935 PE=4 SV=1
A0A1E7R3F5 438 78 13.2 Zonular occludens toxin OS=Acinetobacter proteolyticus OX=1776741 GN=BJD20_20025 PE=4 SV=1
A0A1E1G0N7 712 94 13.2 Toxin secretion ABC transporter ATP-binding and membrane protein OS=Pseudomonas chlororaphis subsp. auran
A0A1D7W733 500 78 13.2 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Brevibacterium linens OX=1703 GN=BLS
A0A1C6IL62 524 78 13.2 Toxin A OS=uncultured Lachnospira sp. OX=446043 GN=toxA PE=4 SV=1
A0A1C6F363 695 92 13.2 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_1 PE=4 SV=1
A0A1C5M4T0 336 78 13.2 Toxin A OS=uncultured Ruminococcus sp. OX=165186 GN=toxA_2 PE=4 SV=1
A0A1B3M3R7 511 78 13.2 Zeta toxin family protein OS=Hydrogenophaga sp. RAC07 OX=1842537 GN=BSY239_529 PE=4 SV=1
A0A1A8GR48 478 78 13.2 Anthrax toxin receptor 2b OS=Nothobranchius korthausae OX=1143690 GN=ANTXR2B PE=4 SV=1
A0A194XQG7 553 78 13.2 Putative HC-toxin efflux carrier OS=Phialocephala scopiformis OX=149040 GN=LY89DRAFT_704077 PE=4 SV=1
A0A194UMU8 534 78 13.2 Putative HC-toxin efflux carrier TOXA OS=Valsa mali var. pyri OX=694573 GN=VP1G_00365 PE=4 SV=1
A0A194UM57 588 78 13.2 Putative HC-toxin efflux carrier TOXA OS=Valsa mali var. pyri OX=694573 GN=VP1G_00052 PE=4 SV=1
A0A176EEU1 434 78 13.2 Toxin HipA OS=Erythrobacter sp. HI0028 OX=1822227 GN=A3723_09595 PE=4 SV=1
A0A175WJP7 564 78 13.2 Putative HC-toxin efflux carrier TOXA OS=Madurella mycetomatis OX=100816 GN=MMYC01_200220 PE=4 SV=1
A0A175VW95 560 78 13.2 Putative HC-toxin efflux carrier TOXA OS=Madurella mycetomatis OX=100816 GN=MMYC01_208638 PE=4 SV=1
A0A174ZUE0 406 78 13.2 Pertussis toxin liberation protein H OS=[Eubacterium] eligens OX=39485 GN=ptlH PE=4 SV=1
A0A167UD15 539 78 13.2 Putative HC-toxin efflux carrier TOXA OS=Penicillium chrysogenum OX=5076 GN=EN45_077440 PE=4 SV=1
A0A166AH36 469 78 13.2 Zonular occludens toxin OS=Vibrio sp. HI00D65 OX=1822216 GN=A3712_08380 PE=4 SV=1
A0A162TTP0 630 83 13.2 Putative toxin component OS=Bacillus subtilis OX=1423 GN=B4122_2875 PE=4 SV=1
A0A162KMV0 549 78 13.2 Bordetella pertussis toxin A OS=Cordyceps confragosa RCEF 1005 OX=1081108 GN=LEL_04101 PE=4 SV=1
A0A109LU77 432 78 13.2 Toxin HipA OS=Erythrobacter sp. AP23 OX=499656 GN=ASS64_13200 PE=4 SV=1
A0A0W8C8E4 599 79 13.2 Multidrug and toxin extrusion protein 1 OS=Phytophthora nicotianae OX=4790 GN=AM587_10006616 PE=4 SV=1
A0A0V7Z7G5 456 78 13.2 Rtx toxin hemolysin-type calcium-binding protein OS=Methylobacterium sp. GXS13 OX=1730094 GN=AO398_13580
A0A0U1MA90 567 78 13.2 Putative HC-toxin efflux carrier TOXA OS=Talaromyces islandicus OX=28573 GN=PISL3812_09515 PE=4 SV=1
A0A0Q9ZJY5 364 78 13.2 Toxin regulator PfoR OS=Psychrobacter sp. P11G3 OX=1699623 GN=AK824_11910 PE=4 SV=1
A0A0Q0K0C5 721 95 13.2 Multidrug/toxin efflux protein, ATP-binding and membrane protein OS=Pseudomonas viridiflava OX=33069 GN=A
A0A0P9ZRW2 721 95 13.2 Multidrug/toxin efflux protein, ATP-binding and membrane protein OS=Pseudomonas syringae pv. ribicola OX=
A0A0P9Y5I4 721 95 13.2 Multidrug/toxin efflux protein, ATP-binding and membrane protein OS=Pseudomonas syringae pv. primulae OX=
A0A0P9QQF7 423 78 13.2 Zonular occludens toxin OS=Pseudomonas ficuserectae OX=53410 GN=ALO69_03076 PE=4 SV=1
A0A0N8S9G3 714 94 13.2 Zonular occludens toxin OS=Pseudomonas amygdali pv. mori OX=34065 GN=ALO63_200106 PE=3 SV=1
A0A0L1IW77 566 78 13.2 MFS toxin efflux pump OS=Aspergillus nomius NRRL 13137 OX=1509407 GN=ANOM_007860 PE=4 SV=1
A0A0K8M4N9 437 78 13.2 Type II toxin-antitoxin system HipA family toxin OS=Pseudomonas syringae pv. actinidiae OX=103796 GN=CUB8
A0A0K1Q9T0 544 78 13.2 RTX toxin OS=Labilithrix luteola OX=1391654 GN=AKJ09_08833 PE=4 SV=1
A0A0K1PV41 514 78 13.2 RTX toxin OS=Labilithrix luteola OX=1391654 GN=AKJ09_04064 PE=4 SV=1
A0A0K1DUK1 437 78 13.2 Toxin HipA OS=Pseudomonas syringae pv. actinidiae ICMP 18884 OX=1095103 GN=IYO_021910 PE=4 SV=1
A0A0H4WY27 509 78 13.2 RTX toxin OS=Myxococcus hansupus OX=1297742 GN=A176_005228 PE=4 SV=1
A0A0F7PML7 636 84 13.2 Ca2+-binding protein, RTX toxin OS=Hoeflea sp. IMCC20628 OX=1620421 GN=IMCC20628_02669 PE=4 SV=1
A0A0E9ESV7 431 78 13.2 Pertussis toxin liberation protein H OS=Chlamydia trachomatis OX=813 GN=ptlH PE=4 SV=1
A0A0A6UPT2 407 78 13.2 Hansenula MRAKII killer toxin-resistant protein 1 OS=Actinoplanes utahensis OX=1869 GN=MB27_18085 PE=4 SV
A0A0A2WJE8 689 91 13.2 Calcium-binding RTX toxin-like protein OS=Lysobacter dokdonensis DS-58 OX=1300345 GN=LF41_837 PE=4 SV=1
A0A093Y878 542 78 13.2 Putative HC-toxin efflux carrier TOXA OS=Talaromyces marneffei PM1 OX=1077442 GN=GQ26_0014360 PE=4 SV=1
A0A093VS73 574 78 13.2 Putative HC-toxin efflux carrier TOXA OS=Talaromyces marneffei PM1 OX=1077442 GN=GQ26_0022550 PE=4 SV=1
A0A090R6M9 475 78 13.2 RTX toxin putative OS=Vibrio sp. C7 OX=1001886 GN=JCM19233_771 PE=4 SV=1
A0A077PLL2 430 78 13.2 C component of insecticidal toxin complex (Tc) OS=Xenorhabdus bovienii str. kraussei Quebec OX=1398203 GN
A0A073CFN5 810 107 13.2 Hemolysin-type calcium-binding toxin OS=Planktothrix agardhii NIVA-CYA 126/8 OX=388467 GN=A19Y_2164 PE=4
A0A063BSC4 416 78 13.2 Toxin biosynthesis protein tri7-like OS=Ustilaginoidea virens OX=1159556 GN=UV8b_5437 PE=4 SV=1
A0A061YPV0 443 78 13.2 Transcriptional regulator OS=Escherichia coli OX=562 GN=AM434_16200 PE=4 SV=1
A0A061PV83 483 78 13.2 Multidrug and toxin extrusion MATE family efflux pump YdhE/NorM OS=Vibrio sp. JCM 19052 OX=1460647 GN=JCM
A0A034T6R9 462 78 13.2 Zona occludens toxin OS=Vibrio sp. JCM 18904 OX=1298599 GN=JCM18904_247 PE=4 SV=1
S4CX79 641 84 13.1 Toxin secretion/phage lysis holin OS=Enterococcus faecalis 06-MB-DW-09 OX=1260358 GN=D922_01280 PE=4 SV=1
Q1I6G1 708 93 13.1 Putative type I toxin efflux ATP-binding membrane protein OS=Pseudomonas entomophila (strain L48) OX=384676 G

N9V9A2 716 94 13.1 Toxin secretion ATP-binding protein OS=Aeromonas diversa CDC 2478-85 OX=1268237 GN=G114_11140 PE=4 SV=1
I4YYX3 741 97 13.1 Ca2+-binding protein, RTX toxin OS=Microvirga lotononidis OX=864069 GN=MicloDRAFT_00016370 PE=4 SV=1
I3J6B0 617 81 13.1 Anthrax toxin receptor 1b OS=Oreochromis niloticus OX=8128 GN=LOC100694769 PE=4 SV=1
I1XJU3 733 96 13.1 RTX toxin transporter OS=Methylophaga nitratireducenticrescens (strain ATCC BAA-2433 / DSM 25689 / JAM1) OX=7
F0ELG2 641 84 13.1 Toxin secretion/phage lysis holin OS=Enterococcus casseliflavus ATCC 12755 OX=888066 GN=HMPREF9087_2254 PE=4
B5FFJ3 628 82 13.1 Toxin corregulated pilus biosynthesis protein I OS=Vibrio fischeri (strain MJ11) OX=388396 GN=VFMJ11_1778 PE=
A0A2I7IAY2 785 103 13.1 RTX toxin OS=Phaeobacter piscinae OX=1580596 GN=PhaeoP71_02414 PE=4 SV=1
A0A2I7GII2 785 103 13.1 RTX toxin OS=Phaeobacter inhibens OX=221822 GN=PhaeoP70_00694 PE=4 SV=1
A0A2I4AT18 647 85 13.1 Multidrug and toxin extrusion protein OS=Austrofundulus limnaeus OX=52670 GN=LOC106514098 PE=3 SV=1
A0A2D2CDS0 627 82 13.1 ADP-ribosylating toxin OS=Staphylococcus epidermidis OX=1282 GN=CPZ21_11340 PE=4 SV=1
A0A2A2DB51 625 82 13.1 Zeta toxin family protein OS=Streptomyces albireticuli OX=1940 GN=CK936_11910 PE=4 SV=1
A0A291FL22 785 103 13.1 RTX toxin OS=Phaeobacter piscinae OX=1580596 GN=PhaeoP36_00626 PE=4 SV=1
A0A267F7B7 633 83 13.1 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig032378g1 PE=3 SV=1
A0A238KS40 731 96 13.1 Toxin RTX-I translocation ATP-binding protein OS=Pelagimonas varians OX=696760 GN=apxIB_1 PE=4 SV=1
A0A222QQK4 686 90 13.1 General secretion pathway protein D (Cholera toxin secretion protein epsD) OS=Escherichia coli NCCP15648
A0A1S3SS83 727 95 13.1 toxin CdiA-like isoform X9 OS=Salmo salar OX=8030 GN=LOC106611486 PE=4 SV=1
A0A1N7LIG3 695 91 13.1 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Zobellia uliginosa OX=143224 GN=SAMN05421
A0A1M6VYP7 616 81 13.1 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Reichenbachiella agariperforans OX=156994
A0A1J9QXV6 664 87 13.1 Protoplast regeneration and killer toxin resistance protein OS=Diplodia corticola OX=236234 GN=BKCO1_3500
A0A1J1LFW4 663 87 13.1 Putative haemolysin-type calcium-binding toxin, RTX-like (Expressed) OS=Planktothrix tepida PCC 9214 OX=6
A0A1ION188 670 88 13.1 Ca2+-binding protein, RTX toxin-related OS=Loktanella koreensis OX=364200 GN=SAMN04488515_0349 PE=4 SV=1
A0A1H9WYW6 826 108 13.1 Ca2+-binding protein, RTX toxin-related OS=Roseivivax roseus OX=641238 GN=SAMN04490244_11519 PE=4 SV=1
A0A1H5VN69 804 105 13.1 Ca2+-binding protein, RTX toxin-related OS=Thalassococcus halodurans OX=373675 GN=SAMN04488045_1114 PE=4
A0A1C6KEM5 712 93 13.1 Toxin A OS=uncultured Clostridium sp. OX=59620 GN=toxA_2 PE=4 SV=1
A0A1C6JJX0 850 111 13.1 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_3 PE=4 SV=1
A0A1C5CWF1 788 103 13.1 Restriction endonuclease fold toxin 5 OS=Streptomyces sp. Ncost-T10-10d OX=1839774 GN=GA0115254_115614 P
A0A1B8RHQ8 680 89 13.1 RTX toxin OS=Rhizobium leguminosarum bv. trifolii OX=386 GN=BAE36_04365 PE=4 SV=1
A0A175WGU0 601 79 13.1 Putative HC-toxin efflux carrier TOXA OS=Madurella mycetomatis OX=100816 GN=MMYC01_200638 PE=4 SV=1
A0A135ZGZ6 610 80 13.1 Cytolysin and hemolysin HlyA Pore-forming toxin OS=Photobacterium sanguinicancri OX=875932 GN=AS132_23675
A0A0T9T8L3 711 93 13.1 RTX family toxin transporter OS=Yersinia kristensenii OX=28152 GN=apxIB_2 PE=4 SV=1
A0A0T9S5S5 708 93 13.1 RTX family toxin transporter OS=Yersinia enterocolitica OX=630 GN=apxIB_1 PE=4 SV=1
A0A0R3MM33 765 100 13.1 RTX toxin OS=Bradyrhizobium retamae OX=1300035 GN=CQ13_34935 PE=4 SV=1
A0A0J8QQ86 725 95 13.1 HC-toxin synthetase OS=Coccidioides immitis RMSCC 3703 OX=454286 GN=CISG_04308 PE=4 SV=1
A0A0B6CRU5 703 92 13.1 Toxin RTX-I translocation ATP-binding protein OS=Francisella philomiragia OX=28110 GN=apxIB PE=4 SV=1
A0A086TGZ4 595 78 13.1 Putative HC-toxin efflux carrier-like protein OS=Acremonium chrysogenum (strain ATCC 11550 / CBS 779.69 /
X7EF43 435 77 13.0 RTX toxin OS=Roseivivax halodurans JCM 10272 OX=1449350 GN=OCH239_06460 PE=4 SV=1
X5U542 435 77 13.0 Toxin HipA OS=Mesorhizobium sp. LSHC426A00 OX=1287298 GN=X762_29530 PE=4 SV=1
V9VY73 434 77 13.0 RTX toxin OS=Leisingera methylohalidivorans DSM 14336 OX=999552 GN=METH_15130 PE=4 SV=1
U9TVF8 464 77 13.0 Toxin biosynthesis protein OS=Rhizophagus irregularis (strain DAOM 181602 / DAOM 197198 / MUCL 43194) OX=7470
T2NRW2 644 84 13.0 Toxin secretion/phage lysis holin OS=Enterococcus faecium 13.SD.W.09 OX=1259824 GN=D931_00985 PE=4 SV=1
R7VCW0 568 77 13.0 Multidrug and toxin extrusion protein OS=Capitella teleta OX=283909 GN=CAPTEDRAFT_214524 PE=3 SV=1
Q9BQC3-3 354 77 13.0 Isoform 3 of 2-(3-amino-3-carboxypropyl)histidine synthase subunit 2 OS=Homo sapiens OX=9606 GN=DPH2
Q93Q17 475 77 13.0 ADP-ribosyltransferase toxin AexT OS=Aeromonas salmonicida OX=645 GN=aexT PE=3 SV=1
Q5I0E9 566 77 13.0 Multidrug and toxin extrusion protein 1 OS=Rattus norvegicus OX=10116 GN=Slc47a1 PE=1 SV=1
Q4SYH4 632 82 13.0 Multidrug and toxin extrusion protein OS=Tetraodon nigroviridis OX=99883 GN=GSTENG00010333001 PE=3 SV=1
Q460N3 678 88 13.0 Poly [ADP-ribose] polymerase 15 OS=Homo sapiens OX=9606 GN=PARP15 PE=1 SV=2
Q2EHL7 457 77 13.0 Toxin and drug export protein A OS=Aggregatibacter actinomycetemcomitans OX=714 GN=tdeA PE=1 SV=1
P40487 425 77 13.0 2-(3-amino-3-carboxypropyl)histidine synthase subunit 1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S28
P39410 443 77 13.0 Toxin YjjJ OS=Escherichia coli (strain K12) OX=83333 GN=yjjJ PE=1 SV=1
N4UAF3 523 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 1) OX=1229664 GN=FOC1
N4U255 528 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 1) OX=1229664 GN=FOC1
N4TSD7 546 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 1) OX=1229664 GN=FOC1
M0CR90 453 77 13.0 Rhizobiocin/RTX toxin and hemolysin-type calcium binding protein OS=Haloterrigena salina JCM 13891 OX=1227488
L0LMC9 466 77 13.0 Multidrug and toxin extrusion (MATE) family efflux pump OS=Rhizobium tropici CIAT 899 OX=698761 GN=RTCIAT899_
K7GIM9 453 77 13.0 Anthrax toxin receptor 2 OS=Pelodiscus sinensis OX=13735 GN=ANTXR2 PE=4 SV=1
K6YGD2 559 77 13.0 Toxin secretion ABC transporter ATP-binding protein OS=Paraglaciecola mesophila KMM 241 OX=1128912 GN=GMES_07
K5VBA9 458 77 13.0 Zonular occludens toxin family protein OS=Vibrio sp. HENC-03 OX=992012 GN=VCHENC03_3030 PE=4 SV=1
J7SKC5 511 77 13.0 Putative toxin subunit OS=Morganella morganii subsp. morganii KT OX=1124991 GN=MU9_1771 PE=4 SV=1
H8NYB5 509 77 13.0 Toxin co-regulated pilus biosynthesis protein T OS=Rahnella aquatilis HX2 OX=1151116 GN=Q7S_20055 PE=4 SV=1
H0EWE5 548 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Glarea lozoyensis (strain ATCC 74030 / MF5533) OX=1104152 GN=M7I_711
H0ESG0 555 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Glarea lozoyensis (strain ATCC 74030 / MF5533) OX=1104152 GN=M7I_564
G8TNJ7 473 77 13.0 Binary exotoxin B/Anthrax toxin B moiety protective antigen OS=Niastella koreensis (strain DSM 17620 / KACC 1
G5DMQ7 652 85 13.0 CryIII crystal toxin protein OS=Bacillus thuringiensis OX=1428 GN=cryIII PE=3 SV=1
G4TBN7 433 77 13.0 Related to MAK11 protein (Maintenance of killer toxin-encoding satellite M1 dsRNA) OS=Serendipita indica (str
F8GY53 614 80 13.0 Toxin corregulated pilus biosynthesis protein T OS=Cupriavidus necator (strain ATCC 43291 / DSM 13513 / N-1)
F7DGR4 486 77 13.0 Anthrax toxin receptor 2 OS=Ornithorhynchus anatinus OX=9258 GN=ANTXR2 PE=4 SV=2
F2I838 513 77 13.0 Putative toxin-antitoxin system toxin component, PIN family OS=Aerococcus urinae (strain ACS-120-V-Col10a) OX
E7EX57 443 77 13.0 Multidrug and toxin extrusion protein 1 OS=Homo sapiens OX=9606 GN=SLC47A1 PE=1 SV=1

E6R723 486 77 13.0 Protoplast regeneration and killer toxin resistance protein, putative OS=Cryptococcus gattii serotype B (stra
E0MSI4 508 77 13.0 Rhizobiocin/RTX toxin and hemolysin-type calcium binding protein OS=Ahrensia sp. R2A130 OX=744979 GN=R2A130_1
C1FUH3 490 77 13.0 Toxin complex component ORF-X3 OS=Clostridium botulinum (strain Kyoto / Type A2) OX=536232 GN=CLM_0891 PE=4 S
A3VHU3 463 77 13.0 Hemolysin-type calcium-binding toxin OS=Maritimibacter alkaliphilus HTCC2654 OX=314271 GN=RB2654_08762 PE=4 S
A1JMA6 708 92 13.0 Putative RTX-family toxin transporter OS=Yersinia enterocolitica serotype O:8 / biotype 1B (strain NCTC 13174
A1CHU3 519 77 13.0 Toxin biosynthesis cytochrome P450 monooxygenase, putative OS=Aspergillus clavatus (strain ATCC 1007 / CBS 51
A1CGF7 487 77 13.0 Toxin biosynthesis cytochrome P450 monooxygenase, putative OS=Aspergillus clavatus (strain ATCC 1007 / CBS 51
A0A2J9TXD4 448 77 13.0 Type II toxin-antitoxin system HipA family toxin OS=Bordetella parapertussis OX=519 GN=AL462_000430 PE=4
A0A2I8XEW1 490 77 13.0 Aerolysin family beta-barrel pore-forming toxin OS=Vibrio campbellii OX=680 GN=C1N51_14980 PE=4 SV=1
A0A2I8VVW0 490 77 13.0 Aerolysin family beta-barrel pore-forming toxin OS=Vibrio campbellii OX=680 GN=C1N50_07760 PE=4 SV=1
A0A2I7M0R8 785 102 13.0 RTX toxin OS=Phaeobacter inhibens OX=221822 GN=PhaeoP66_02655 PE=4 SV=1
A0A2I7K6C3 785 102 13.0 RTX toxin OS=Phaeobacter inhibens OX=221822 GN=PhaeoP88_00755 PE=4 SV=1
A0A2I7GVL3 785 102 13.0 RTX toxin OS=Phaeobacter inhibens OX=221822 GN=PhaeoP30_00687 PE=4 SV=1
A0A2I4C1X8 404 77 13.0 multidrug and toxin extrusion protein 1-like OS=Austrofundulus limnaeus OX=52670 GN=LOC106524647 PE=4 SV=
A0A2I1DFK2 436 77 13.0 Toxin biosynthesis protein OS=Aspergillus campestris IBT 28561 OX=1392248 GN=P168DRAFT_278213 PE=4 SV=1
A0A2I1CB52 583 77 13.0 Putative MFS toxin transporter OS=Aspergillus novofumigatus IBT 16806 OX=1392255 GN=P174DRAFT_458469 PE=4
A0A2H3Q366 515 77 13.0 Mosquitocidal toxin protein OS=Bacillus sp. AFS012607 OX=2033485 GN=CN409_27100 PE=4 SV=1
A0A2G7F6Z9 747 97 13.0 SUKH-4 immunity protein of toxin-antitoxin system OS=Streptomyces sp. 70 OX=1827606 GN=CLW09_07162 PE=4 S
A0A2G7BHV0 448 77 13.0 Antitoxin component YwqK of YwqJK toxin-antitoxin module OS=Janthinobacterium sp. 35 OX=2035210 GN=CLU93_
A0A2G6YJ54 729 95 13.0 SUKH-4 immunity protein of toxin-antitoxin system OS=Streptomyces sp. 76 OX=2035220 GN=CLW13_1429 PE=4 SV
A0A2G5I1X3 552 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Cercospora beticola OX=122368 GN=CB0940_02347 PE=4 SV=1
A0A2G4DJ07 440 77 13.0 Toxin HipA OS=Pseudomonas sp. NZIPFR-PS5 OX=1230465 GN=AO263_33635 PE=4 SV=1
A0A2G1ARW2 490 77 13.0 Aerolysin family beta-barrel pore-forming toxin OS=Vibrio splendidus OX=29497 GN=CSB62_18805 PE=4 SV=1
A0A2E3J967 420 77 13.0 Toxin HipA OS=Gemmatimonadetes bacterium OX=2026742 GN=CME25_24305 PE=4 SV=1
A0A2D9UUZ8 434 77 13.0 RTX toxin OS=Thioclava sp. OX=1933450 GN=CMO21_06100 PE=4 SV=1
A0A2D8DI48 454 77 13.0 Type II toxin-antitoxin system HipA family toxin OS=Thalassospira sp. OX=1912094 GN=CMO03_12610 PE=4 SV=1
A0A2B8ZLN6 515 77 13.0 Mosquitocidal toxin protein OS=Bacillus thuringiensis OX=1428 GN=COJ78_31215 PE=4 SV=1
A0A2A8PB11 516 77 13.0 Mosquitocidal toxin protein OS=Bacillus cereus OX=1396 GN=CN419_22390 PE=4 SV=1
A0A2A6HRD3 687 89 13.0 RTX toxin OS=Rhizobium sp. L43 OX=2035452 GN=CO667_13600 PE=4 SV=1
A0A267GSQ3 568 77 13.0 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig017756g2 PE=3 SV=1
A0A267FYC8 574 77 13.0 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig026985g3 PE=3 SV=1
A0A267FPJ1 633 82 13.0 Multidrug and toxin extrusion protein OS=Macrostomum lignano OX=282301 GN=BOX15_Mlig032378g3 PE=3 SV=1
A0A259KFV3 440 77 13.0 Toxin HipA OS=Polaromonas sp. 39-63-203 OX=1970419 GN=B7X59_05175 PE=4 SV=1
A0A259B3I6 440 77 13.0 Toxin HipA OS=Polaromonas sp. 24-62-144 OX=1970414 GN=B7Y03_06230 PE=4 SV=1
A0A258TXV3 440 77 13.0 Toxin HipA OS=Polaromonas sp. 28-63-22 OX=1970416 GN=B7Y42_09525 PE=4 SV=1
A0A258QE69 440 77 13.0 Toxin HipA OS=Polaromonas sp. 35-63-240 OX=1970417 GN=B7Y54_05665 PE=4 SV=1
A0A258CD43 441 77 13.0 Toxin HipA OS=Caulobacterales bacterium 32-67-6 OX=1970502 GN=B7Z13_07165 PE=4 SV=1
A0A249PSU8 466 77 13.0 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Sinorhizobium fredii CCBAU 83666 OX=
A0A246U2E9 648 84 13.0 Toxin OS=Rhizobium sp. R693 OX=1764276 GN=ATY79_02040 PE=4 SV=1
A0A244DY26 449 77 13.0 Toxin HipA OS=Paraburkholderia terrae OX=311230 GN=CA603_05925 PE=4 SV=1
A0A239W7R9 485 77 13.0 Pertussis toxin liberation protein H OS=Cutibacterium granulosum OX=33011 GN=ptlH PE=4 SV=1
A0A239G4W7 437 77 13.0 Ca2+-binding protein, RTX toxin-related OS=Rhodobacter megalophilus OX=418630 GN=SAMN05421763_102551 PE=4
A0A238KB20 715 93 13.0 Toxin RTX-I translocation ATP-binding protein OS=Ruegeria arenilitoris OX=1173585 GN=apxIB PE=4 SV=1
A0A212QHL7 622 81 13.0 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Marinobacter sp. es.042 OX=176
A0A210QXP8 593 77 13.0 Multidrug and toxin extrusion protein OS=Mizuhopecten yessoensis OX=6573 GN=KP79_PYT13631 PE=3 SV=1
A0A1Y2TRA7 440 77 13.0 Putative toxin biosynthesis protein OS=Hypoxylon sp. EC38 OX=1001937 GN=K449DRAFT_465177 PE=4 SV=1
A0A1W9XRP5 434 77 13.0 RTX toxin OS=Thioclava sp. E1Ox9 OX=1549850 GN=AKL02_06840 PE=4 SV=1
A0A1V5EJK6 414 77 13.0 Toxin HipA OS=Bradyrhizobium sp. BR10280 OX=1399419 GN=A5906_08220 PE=4 SV=1
A0A1T2BD81 434 77 13.0 RTX toxin OS=Thioclava sp. DLFJ5-1 OX=1915314 GN=BMI86_08085 PE=4 SV=1
A0A1T2A4J1 434 77 13.0 RTX toxin OS=Thioclava sp. F28-4 OX=1915315 GN=BMI87_09370 PE=4 SV=1
A0A1S3SSB0 669 87 13.0 toxin CdiA-like isoform X15 OS=Salmo salar OX=8030 GN=LOC106611486 PE=4 SV=1
A0A1S3SS97 493 77 13.0 toxin CdiA-like isoform X17 OS=Salmo salar OX=8030 GN=LOC106611486 PE=4 SV=1
A0A1S3RQA3 365 77 13.0 multidrug and toxin extrusion protein 1-like OS=Salmo salar OX=8030 GN=LOC106604060 PE=4 SV=1
A0A1S2P140 584 77 13.0 Zeta toxin family protein OS=Streptomyces sp. MUSC 14 OX=1354889 GN=BIV25_39115 PE=4 SV=1
A0A1R3F8A7 474 77 13.0 Zonular occludens toxin OS=Vibrio sp. 10N.222.47.A9 OX=1903178 GN=BH582_14295 PE=4 SV=1
A0A1Q8T5A1 448 77 13.0 Toxin HipA OS=Salinicola sp. MH3R3-1 OX=1928762 GN=BTW08_05935 PE=4 SV=1
A0A1Q8S6V5 585 77 13.0 Putative HC-toxin efflux carrier TOXA 7 OS=Colletotrichum chlorophyti OX=708187 GN=CCHL11_02025 PE=4 SV=1
A0A1N7NUU1 474 77 13.0 Ca2+-binding protein, RTX toxin-related OS=Gemmobacter megaterium OX=1086013 GN=SAMN05421774_104147 PE=4
A0A1N6XIZ3 569 77 13.0 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Pseudomonas sp. B10 OX=118613
A0A1M8H718 404 77 13.0 Cholera toxin secretion protein epsF OS=Mycobacterium abscessus subsp. abscessus OX=1185650 GN=epsF_1 PE=
A0A1M7DL29 513 77 13.0 Restriction endonuclease fold toxin 5 OS=Chryseobacterium carnipullorum OX=1124835 GN=SAMN05444360_104235
A0A1M6S6R9 428 77 13.0 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Fibrobacter sp. UWH4 OX=189621
A0A1M6R590 440 77 13.0 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Reichenbachiella agariperforans OX=156994
A0A1M5WUT2 428 77 13.0 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Fibrobacter sp. UWCM OX=189620
A0A1M4YXA5 515 77 13.0 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Chryseobacterium sp. YR203 OX=
A0A1K2EBK6 440 77 13.0 Nucleotidyl transferase AbiEii toxin, Type IV TA system OS=Streptomyces atratus OX=1893 GN=SAMN02787144_1
A0A1I6IWU7 622 81 13.0 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=[Clostridium] aminophilum OX=1
A0A1I4I5D9 714 93 13.0 Ca2+-binding protein, RTX toxin-related OS=Methylobacterium salsuginis OX=414703 GN=SAMN04488125_11666 PE

A0A1I3QBU2 691 90 13.0 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Caulobacter sp. UNC279MFTsu5.1
A0A1I1B8V4 441 77 13.0 Ca2+-binding protein, RTX toxin-related OS=Nocardioides alpinus OX=748909 GN=SAMN05192575_11537 PE=4 SV=1
A0A1IOLKP8 416 77 13.0 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Myxococcus fulvus OX=33 GN=SAM
A0A1H8B602 746 97 13.0 Ca2+-binding protein, RTX toxin-related OS=Gemmobacter aquatilis OX=933059 GN=SAMN04488103_10267 PE=4 SV=
A0A1H7IGG9 439 77 13.0 Nucleotidyl transferase AbiEii toxin, Type IV TA system OS=Streptacidiphilus jiangxiensis OX=235985 GN=SA
A0A1H6YFA6 527 77 13.0 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Cyclobacterium halophilum OX=1416801 GN=S
A0A1H3PY81 721 94 13.0 Ca2+-binding protein, RTX toxin-related OS=Nitrosomonas sp. Nm58 OX=200126 GN=SAMN05421754_104913 PE=4 SV
A0A1G8T264 593 77 13.0 Ca2+-binding protein, RTX toxin-related OS=Citreicella marina OX=555512 GN=SAMN04487993_102760 PE=4 SV=1
A0A1G7Z894 558 77 13.0 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Propionivibrio dicarboxylicus
A0A1G3GJK1 435 77 13.0 RTX toxin OS=Rhodobacterales bacterium RIFCSPHIGHO2_02_FULL_62_130 OX=1802014 GN=A3D16_00080 PE=4 SV=1
A0A1G3GB29 428 77 13.0 RTX toxin OS=Rhodobacterales bacterium RIFCSPHIGHO2_02_FULL_62_130 OX=1802014 GN=A3D16_02300 PE=4 SV=1
A0A1G0R662 447 77 13.0 Toxin HipA OS=Hydrogenophilales bacterium RIFOXYA1_FULL_63_33 OX=1798414 GN=A2199_06380 PE=4 SV=1
A0A1F8A2C4 561 77 13.0 MFS toxin efflux pump OS=Aspergillus bombycis OX=109264 GN=ABOM_006000 PE=4 SV=1
A0A1F3NNM9 424 77 13.0 Toxin HipA OS=Bacteroidetes bacterium RBG_13_42_15 OX=1797355 GN=A2Y71_05055 PE=4 SV=1
A0A1E5EQ16 474 77 13.0 Zonular occludens toxin OS=Vibrio splendidus 1F-157 OX=617145 GN=A148_24345 PE=4 SV=1
A0A1C6L278 553 77 13.0 RTX-I toxin determinant B OS=uncultured Bacteroides sp. OX=162156 GN=apxIB PE=4 SV=1
A0A1C6KAS5 475 77 13.0 Toxin A OS=uncultured Ruminococcus sp. OX=165186 GN=toxA_3 PE=4 SV=1
A0A1C5UZC7 453 77 13.0 Toxin A OS=uncultured Blautia sp. OX=765821 GN=toxA_4 PE=4 SV=1
A0A1C5TR05 625 81 13.0 Toxin B OS=uncultured Clostridium sp. OX=59620 GN=toxB_1 PE=4 SV=1
A0A1C5MH01 475 77 13.0 Toxin A OS=uncultured Ruminococcus sp. OX=165186 GN=toxA_2 PE=4 SV=1
A0A1C5KM64 522 77 13.0 Toxin A OS=uncultured Clostridium sp. OX=59620 GN=toxA_6 PE=4 SV=1
A0A1C5E757 421 77 13.0 Nucleotidyl transferase AbiEii toxin, Type IV TA system OS=Streptomyces sp. Ncost-T10-10d OX=1839774 GN=G
A0A1C4F9X6 414 77 13.0 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Rhodococcus enclensis OX=10495
A0A1C1CS02 602 78 13.0 Putative HC-toxin efflux carrier TOXA OS=Cladophialophora carrionii OX=86049 GN=TOXA PE=4 SV=1
A0A1C1CLN7 550 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Cladophialophora carrionii OX=86049 GN=TOXA PE=4 SV=1
A0A1COU3Q3 504 77 13.0 Toxin coregulated pilus biosynthesis protein T OS=Photorhabdus asymbiotica subsp. australis OX=286156 GN=
A0A1B9QE70 474 77 13.0 Zonular occludens toxin OS=Vibrio splendidus OX=29497 GN=A6D94_19305 PE=4 SV=1
A0A1B8QBC3 449 77 13.0 Toxin HipA OS=Moraxella atlantae OX=34059 GN=A9308_07370 PE=4 SV=1
A0A1B1PWE7 436 77 13.0 RTX toxin OS=Yangia sp. CCB-MM3 OX=1792508 GN=AYJ57_20660 PE=4 SV=1
A0A1A8UX06 583 77 13.0 Anthrax toxin receptor 2a OS=Nothobranchius furzeri OX=105023 GN=ANTXR2A PE=4 SV=1
A0A1A8E1W6 478 77 13.0 Anthrax toxin receptor 2b OS=Nothobranchius kadleci OX=1051664 GN=ANTXR2B PE=4 SV=1
A0A1A8B114 646 84 13.0 Multidrug and toxin extrusion protein OS=Nothobranchius furzeri OX=105023 GN=SLC47A1 PE=3 SV=1
A0A1A7XWB7 501 77 13.0 Anthrax toxin receptor 2a OS=Aphyosemion striatum OX=60296 GN=ANTXR2A PE=4 SV=1
A0A1A7UYH1 714 93 13.0 Toxin RTX-I translocation ATP-binding protein OS=Vibrio mediterranei OX=689 GN=apxIB_4 PE=4 SV=1
A0A194WOC7 560 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Valsa mali OX=105487 GN=VM1G_04964 PE=4 SV=1
A0A194V711 587 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Valsa mali var. pyri OX=694573 GN=VP1G_06933 PE=4 SV=1
A0A178ATB1 572 77 13.0 Putative HC-toxin efflux carrier OS=Stagonospora sp. SRC1lsM3a OX=765868 GN=IQ06DRAFT_326418 PE=4 SV=1
A0A177NVJ0 420 77 13.0 Toxin HipA OS=Methylomonas koyamae OX=702114 GN=A1356_19745 PE=4 SV=1
A0A175W515 557 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Madurella mycetomatis OX=100816 GN=MMYC01_202572 PE=4 SV=1
A0A174NSZ0 517 77 13.0 Toxin A OS=Blautia wexlerae OX=418240 GN=toxA_3 PE=4 SV=1
A0A173QUP3 424 77 13.0 Toxin A OS=Blautia hydrogenotrophica OX=53443 GN=toxA_1 PE=4 SV=1
A0A167V2D2 536 77 13.0 Putative HC-toxin efflux carrier OS=Penicillium chrysogenum OX=5076 GN=EN45_000400 PE=4 SV=1
A0A161IJA5 575 77 13.0 Toxin RTX-I translocation ATP-binding protein OS=Isoptericola dokdonensis DS-3 OX=1300344 GN=apxIB PE=4 S
A0A16OFQB0 455 77 13.0 Toxin HipA OS=Burkholderia sp. OLGA172 OX=1804984 GN=AYM40_21920 PE=4 SV=1
A0A159Z6Y6 490 77 13.0 RTX toxin OS=Defluviimonas alba OX=1335048 GN=AKL17_3924 PE=4 SV=1
A0A158JI25 457 77 13.0 Toxin-related secretion protein OS=Caballeronia arvi OX=1777135 GN=AWB74_03858 PE=4 SV=1
A0A157ZTK5 457 77 13.0 Toxin-related secretion protein OS=Caballeronia catudaia OX=1777136 GN=AWB75_01189 PE=4 SV=1
A0A108WGT6 421 77 13.0 Zonular occludens toxin OS=Pseudomonas amygdali pv. eriobotryae OX=129137 GN=AL052_21085 PE=4 SV=1
A0A0W0SSP5 701 91 13.0 Toxin secretion ATP binding protein OS=Legionella brunensis OX=29422 GN=Lbru_0364 PE=4 SV=1
A0A0U1M715 599 78 13.0 Putative HC-toxin efflux carrier TOXA OS=Talaromyces islandicus OX=28573 GN=PISL3812_08376 PE=4 SV=1
A0A0U1LWW6 524 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Talaromyces islandicus OX=28573 GN=PISL3812_04914 PE=4 SV=1
A0A0U1HFN3 708 92 13.0 RTX family toxin transporter OS=Yersinia enterocolitica OX=630 GN=apxIB_1 PE=4 SV=1
A0A0T9J6Z6 638 83 13.0 Toxin ABC transporter ATP-binding protein/permease OS=Yersinia pseudotuberculosis OX=633 GN=apxIB_2 PE=4
A0A0R4ICZ6 478 77 13.0 Anthrax toxin receptor 2a OS=Danio rerio OX=7955 GN=antxr2a PE=4 SV=1
A0A0R3MAZ8 414 77 13.0 Toxin HipA OS=Bradyrhizobium retamae OX=1300035 GN=CQ13_36375 PE=4 SV=1
A0A0Q4T4S5 444 77 13.0 Toxin HipA OS=Pseudomonas sp. Leaf83 OX=1736239 GN=ASF15_18485 PE=4 SV=1
A0A0N1H1U4 545 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Phialophora attae OX=1664694 GN=AB675_1093 PE=4 SV=1
A0A0M7HV55 520 77 13.0 Toxin co-regulated pilus biosynthesis protein Q OS=Achromobacter sp. OX=134375 GN=ERS370011_03964 PE=4 SV
A0A0LONOP3 565 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Tolypocladium ophioglossoides CBS 100239 OX=1163406 GN=TOPH_0771
A0A0K8QML0 754 98 13.0 Ca2+-binding protein, RTX toxin OS=Mizugakiibacter sediminis OX=1475481 GN=MBSD_2207 PE=4 SV=1
A0A0K8L568 563 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Aspergillus udagawae OX=91492 GN=AUD_2157 PE=4 SV=1
A0A0K1QGM6 493 77 13.0 RTX toxin OS=Labilithrix luteola OX=1391654 GN=AKJ09_11465 PE=4 SV=1
A0A0J6VIJ6 520 77 13.0 Zeta toxin OS=Mycobacterium chlorophenolicum OX=37916 GN=MCHLDSM_05718 PE=4 SV=1
A0A0H6WMT1 518 77 13.0 RTX toxin RtxA OS=Vibrio cholerae OX=666 GN=rtxA_2 PE=4 SV=1
A0A0F8BD86 547 77 13.0 Multidrug and toxin extrusion protein OS=Larimichthys crocea OX=215358 GN=EH28_09896 PE=3 SV=1
A0A0F3FVG2 491 77 13.0 Toxin OS=Clostridium baratii OX=1561 GN=UC77_04375 PE=4 SV=1
A0A0F2D6W4 709 92 13.0 Toxin expression transcriptional accessory protein OS=Streptococcus oralis subsp. oralis OX=1891914 GN=te
A0A0D6DVN8 525 77 13.0 Putative LXG domain-containing toxin OS=Lactococcus piscium MKFS47 OX=297352 GN=LACPI_0355 PE=4 SV=1

A0A0C2F1G6 478 77 13.0 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Pseudomonas batumici OX=226910 GN=UC
A0A0C2D7A0 518 77 13.0 RTX toxin protein OS=Enhygromyxa salina OX=215803 GN=DB30_05098 PE=4 SV=1
A0A0A2VK33 569 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Beauveria bassiana D1-5 OX=1245745 GN=BBAD15_g6456 PE=4 SV=1
A0A094T324 428 77 13.0 Pre-toxin domain with VENN motif family protein OS=Yersinia frederiksenii ATCC 33641 OX=349966 GN=DJ58_44
A0A093V385 731 95 13.0 Putative HC-toxin efflux carrier TOXA OS=Talaromyces marneffei PM1 OX=1077442 GN=GQ26_0181380 PE=4 SV=1
A0A093URC9 517 77 13.0 Putative HC-toxin efflux carrier TOXA OS=Talaromyces marneffei PM1 OX=1077442 GN=GQ26_0440400 PE=4 SV=1
A0A090M5N7 435 77 13.0 Metridin-like ShK toxin OS=Ostreococcus tauri OX=70448 GN=OT_ostta04g00125 PE=4 SV=1
A0A087S7Q8 523 77 13.0 RTX-III toxin determinant A from serotype 8 protein OS=Marine Group I thaumarchaeote SCGC AAA799-B03 OX=1
A0A087RTR5 523 77 13.0 RTX-III toxin determinant A from serotype 8 protein OS=Marine Group I thaumarchaeote SCGC AAA799-D11 OX=1
A0A085V201 430 77 13.0 Toxin HipA OS=Pseudomonas syringae OX=317 GN=IV02_19185 PE=4 SV=1
A0A084WQJ0 544 77 13.0 Alpha-glucosidase binding-toxin receptor OS=Anopheles sinensis OX=74873 GN=ZHAS_00020727 PE=4 SV=1
A0A081S7M8 516 77 13.0 RTX-III toxin determinant A from serotype 8 protein OS=Marine Group I thaumarchaeote SCGC AAA799-E16 OX=1
A0A078AE64 484 77 13.0 Multidrug and toxin extrusion protein 2-like OS=Stylonychia lemnae OX=5949 GN=Contig8496.g9066 PE=4 SV=1
A0A077FET3 718 93 13.0 Multidrug/toxin efflux protein, ATP-binding and membrane protein OS=Pseudomonas alkylphenolica OX=237609
A0A061PEY6 402 77 13.0 Multidrug and toxin extrusion MATE family efflux pump YdhE/NorM OS=Geomicrobium sp. JCM 19039 OX=1460636
A0A059PY44 491 77 13.0 Toxin complex component ORF-X3 OS=Clostridium baratii OX=1561 PE=4 SV=1
A0A059LO25 396 77 13.0 Toxin OS=Pseudomonas mandelii PD30 OX=1419583 GN=V466_19030 PE=4 SV=1
A0A022FRU8 451 77 13.0 Toxin HipA OS=Cupriavidus sp. SK-4 OX=574750 GN=CF68_10555 PE=4 SV=1
X8JV33 514 76 12.9 Protoplast regeneration and killer toxin resistance protein OS=Rhizoctonia solani AG-3 Rhs1AP OX=1086054 GN=R
W4Q3P7 476 76 12.9 Multidrug and toxin extrusion OS=Bacillus wakoensis JCM 9140 OX=1236970 GN=JCM9140_2652 PE=3 SV=1
V2J499 431 76 12.9 Toxin HipA OS=Cupriavidus sp. HPC(L) OX=1217418 GN=B551_0207430 PE=4 SV=1
S9QNQ1 449 76 12.9 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Litoreibacter arenae DSM 19593 OX=112336
S2QW56 382 76 12.9 Membrane protein, putative toxin regulator OS=Lactobacillus paracasei subsp. paracasei Lpp14 OX=1256204 GN=Lp
R4ITQ2 504 76 12.9 Zeta toxin family protein OS=Pseudomonas migulae OX=78543 GN=pD2RT_023 PE=4 SV=1
Q8KHU9 626 81 12.9 HA-70 OS=Clostridium botulinum OX=1491 GN=ha70 PE=1 SV=1
Q75AZ9 426 76 12.9 2-(3-amino-3-carboxypropyl)histidine synthase subunit 1 OS=Ashbya gossypii (strain ATCC 10895 / CBS 109.51 /
Q4FQB6 364 76 12.9 Probable toxin regulator pfoR OS=Psychrobacter arcticus (strain DSM 17307 / 273-4) OX=259536 GN=pfoR PE=4 SV=
Q45882 613 79 12.9 Pesticidal crystal-like protein Cry16Aa OS=Paraclostridium bifermentans OX=1490 GN=cry16Aa PE=1 SV=1
Q19VG7 478 76 12.9 Anthrax toxin receptor 2a OS=Danio rerio OX=7955 GN=antxr2a PE=2 SV=1
N4USX7 479 76 12.9 Putative HC-toxin efflux carrier TOXA OS=Fusarium oxysporum f. sp. cubense (strain race 1) OX=1229664 GN=FOC1
M4V3H9 455 76 12.9 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM,-like protein OS=Ralstonia solanacearum FQY
M3XAJ3 532 76 12.9 Multidrug and toxin extrusion protein OS=Felis catus OX=9685 PE=3 SV=2
K1XSR0 713 92 12.9 Protoplast regeneration and killer toxin resistance protein OS=Marssonina brunnea f. sp. multigermtubi (strai
K0G649 508 76 12.9 Toxin OS=Bacillus thuringiensis MC28 OX=1195464 GN=MC28_E051 PE=4 SV=1
J5MVB1 563 76 12.9 Putative toxin secretion ATP-binding ABC transporter protein OS=Rhizobium sp. CCGE 510 OX=1132836 GN=RCCGE510
I8U5U1 415 76 12.9 Toxin biosynthesis protein (Tri7), putative OS=Aspergillus oryzae (strain 3.042) OX=1160506 GN=Ao3042_00532 P
H9CJF9 486 76 12.9 Toxin co-regulated pilus biosynthesis protein C, outer membrane protein OS=Vibrio cholerae O37 OX=185332 PE=4
H3BXX3 487 76 12.9 Anthrax toxin receptor 2a OS=Tetraodon nigroviridis OX=99883 PE=4 SV=1
H2ZWV6 494 76 12.9 Multidrug and toxin extrusion protein OS=Latimeria chalumnae OX=7897 PE=3 SV=1
G3RJF3 586 76 12.9 Multidrug and toxin extrusion protein OS=Gorilla gorilla gorilla OX=9595 GN=SLC47A1 PE=3 SV=1
G0WPM4 504 76 12.9 Zeta toxin family protein OS=Pseudomonas putida (strain DOT-T1E) OX=1196325 PE=4 SV=1
G0AGW5 371 76 12.9 Zonula occludens toxin-like protein OS=Collimonas fungivorans (strain Ter331) OX=1005048 GN=CFU_2140 PE=4 SV=
F6VZK0 618 80 12.9 Multidrug and toxin extrusion protein OS=Ornithorhynchus anatinus OX=9258 GN=SLC47A2 PE=3 SV=2
E9UX89 414 76 12.9 Toxin-antitoxin system, toxin component, HipA family OS=Nocardioidaceae bacterium Broad-1 OX=408672 GN=NBCG_0
E0WUJ8 432 76 12.9 Putative RTX-like toxin OS=Candidatus Regiella insecticola LSR1 OX=663321 GN=REG_1790 PE=4 SV=1
D6E6K1 354 76 12.9 Predicted membrane protein, putative toxin regulator OS=Gordonibacter pamelaeae 7-10-1-b OX=657308 GN=GPA_023
D2YK12 620 80 12.9 Toxin coregulated pilus biosynthesis protein I OS=Vibrio mimicus VM573 OX=671076 GN=VMD_00960 PE=4 SV=1
D2Q5W2 477 76 12.9 Zeta toxin family protein OS=Bifidobacterium dentium (strain ATCC 27534 / DSM 20436 / JCM 1195 / Bd1) OX=4014
D0Z4Z3 514 76 12.9 ADP-ribosyltransferase toxin aexT (Exoenzyme T) OS=Photobacterium damselae subsp. damselae CIP 102761 OX=6758
B8MXK8 415 76 12.9 Toxin biosynthesis protein (Tri7), putative OS=Aspergillus flavus (strain ATCC 200026 / FGSC A1120 / NRRL 335
A7SLX5 465 76 12.9 2-(3-amino-3-carboxypropyl)histidine synthase subunit 1 OS=Nematostella vectensis OX=45351 GN=dph1 PE=3 SV=1
A7GBF7 490 76 12.9 Toxin complex component ORF-X3 OS=Clostridium botulinum (strain Langeland / NCTC 10281 / Type F) OX=441772 GN
A4QP34 478 76 12.9 Anthrax toxin receptor 2a OS=Danio rerio OX=7955 GN=antxr2a PE=2 SV=1
A0A2K3QHQ1 637 82 12.9 Killer toxin subunits alpha/beta OS=Tolypocladium capitatum OX=45235 GN=TCAP_02993 PE=4 SV=1
A0A2K2U2N7 428 76 12.9 Type II toxin-antitoxin system HipA family toxin OS=Eggerthellaceae bacterium ResAG-85 OX=2070688 GN=C2L8
A0A2J9Y7B0 442 76 12.9 Type II toxin-antitoxin system HipA family toxin OS=Vibrio vulnificus OX=672 GN=AL547_019250 PE=4 SV=1
A0A2J9SI84 449 76 12.9 Type II toxin-antitoxin system HipA family toxin OS=Yersinia enterocolitica OX=630 GN=A6J66_022785 PE=4 S
A0A2J9KQQ4 417 76 12.9 Type II toxin-antitoxin system HipA family toxin OS=Mobiluncus mulieris OX=2052 GN=CEP82_007310 PE=4 SV=1
A0A2J6V4G4 474 76 12.9 Zonular occludens toxin OS=Vibrio cyclitrophicus OX=47951 GN=BCU60_11895 PE=4 SV=1
A0A2I7KJA8 785 101 12.9 RTX toxin OS=Phaeobacter inhibens OX=221822 GN=PhaeoP72_00683 PE=4 SV=1
A0A2I7HUR3 785 101 12.9 RTX toxin OS=Phaeobacter inhibens OX=221822 GN=PhaeoP54_00674 PE=4 SV=1
A0A2I7HNG1 785 101 12.9 RTX toxin OS=Phaeobacter inhibens OX=221822 GN=PhaeoP78_02546 PE=4 SV=1
A0A2I6JNZ1 443 76 12.9 Transcriptional regulator OS=Escherichia coli OX=562 GN=CR916_03640 PE=4 SV=1
A0A2I2YIR5 565 76 12.9 Anthrax toxin receptor like OS=Gorilla gorilla gorilla OX=9595 GN=ANTXRL PE=4 SV=1
A0A2I0GH95 364 76 12.9 Toxin regulator PfoR OS=Psychrobacter sp. Choline-02u-9 OX=2058310 GN=CXF69_04705 PE=4 SV=1
A0A2I0EDE9 552 76 12.9 Toxin ABC transporter OS=Shewanella sp. GutCb OX=2058315 GN=CXF86_13080 PE=4 SV=1
A0A2I0E5Y4 364 76 12.9 Toxin regulator PfoR OS=Psychrobacter sp. Choline-02u-13 OX=2058308 GN=CXF56_03410 PE=4 SV=1
A0A2H9U469 443 76 12.9 Toxin OS=Aeromonas cavernicola OX=1006623 GN=CUC53_10390 PE=4 SV=1

A0A2G7FJF6 426 76 12.9 Toxin biosynthesis protein (Tri7) OS=Aspergillus arachidicola OX=656916 GN=AARAC_006535 PE=4 SV=1
A0A2G6Z2T1 443 76 12.9 Nucleic acid/nucleotide deaminase of polymorphic system toxin OS=Streptomyces sp. 2233.5 OX=1938839 GN=BX
A0A2G2A2C6 432 76 12.9 RTX toxin OS=Thalassobium sp. OX=2030825 GN=COB65_09265 PE=4 SV=1
A0A2G1APQ0 474 76 12.9 Zonular occludens toxin OS=Vibrio splendidus OX=29497 GN=CSB62_23295 PE=4 SV=1
A0A2E2JVZ1 435 76 12.9 RTX toxin OS=Rhodovulum sp. OX=34009 GN=CMM86_12885 PE=4 SV=1
A0A2D8Q6M7 440 76 12.9 Toxin HipA OS=Tistrella sp. OX=2024861 GN=CMO29_06135 PE=4 SV=1
A0A2D8C2D8 433 76 12.9 Toxin HipA OS=Algoriphagus sp. OX=1872435 GN=CL554_03610 PE=4 SV=1
A0A2D5BZU5 449 76 12.9 Type II toxin-antitoxin system HipA family toxin OS=Alteromonas sp. OX=232 GN=CL593_06075 PE=4 SV=1
A0A2D4ZHG1 414 76 12.9 Toxin secretion protein OS=Pseudoalteromonas sp. OX=53249 GN=CMK67_04355 PE=4 SV=1
A0A2D0SY49 568 76 12.9 Multidrug and toxin extrusion protein OS=Ictalurus punctatus OX=7998 GN=LOC108278408 PE=3 SV=1
A0A2C9A5S8 442 76 12.9 Antitoxin component YwqK of the YwqJK toxin-antitoxin module OS=Fibrobacter sp. UWT3 OX=1896225 GN=SAMN05
A0A2C8XF83 568 76 12.9 Papain fold toxin 1, glutamine deamidase OS=Streptomyces sp. OK228 OX=1882786 GN=SAMN05442782_7816 PE=4 S
A0A2C5UCJ6 429 76 12.9 Type II toxin-antitoxin system HipA family toxin OS=Klebsiella oxytoca OX=571 GN=CRX54_08540 PE=4 SV=1
A0A2C2DPG1 516 76 12.9 Mosquitocidal toxin protein OS=Bacillus cereus OX=1396 GN=COA01_27595 PE=4 SV=1
A0A2B9SN68 590 76 12.9 Toxin OS=Bacillus cereus OX=1396 GN=CN980_23780 PE=4 SV=1
A0A2A8HX45 431 76 12.9 Toxin HipA OS=Novosphingobium sp. PC22D OX=1962403 GN=B2G71_04040 PE=4 SV=1
A0A2A3IR89 443 76 12.9 Nucleic acid/nucleotide deaminase of polymorphic system toxin OS=Streptomyces sp. 2321.6 OX=1938840 GN=BX
A0A291FXF4 785 101 12.9 RTX toxin OS=Phaeobacter piscinae OX=1580596 GN=PhaeoP13_00610 PE=4 SV=1
A0A286XBE8 564 76 12.9 Anthrax toxin receptor 1 OS=Cavia porcellus OX=10141 GN=ANTXR1 PE=4 SV=1
A0A286EPE3 464 76 12.9 Putative RNase toxin 24 of polymorphic toxin system OS=Jatrophihabitans sp. GAS493 OX=1907575 GN=SAMN0589
A0A285J8N8 658 85 12.9 Predicted ribonuclease, toxin component of the YeeF-YezG toxin-antitoxin module OS=Bacillus sp. GL120224-
A0A261F120 471 76 12.9 Papain fold toxin 1 OS=Pseudoscardovia suis OX=987063 GN=PSSU_0393 PE=4 SV=1
A0A259MUZ3 438 76 12.9 Toxin HipA OS=Rhodospirillales bacterium 39-66-50 OX=1970570 GN=B7X63_02785 PE=4 SV=1
A0A246HKJ6 435 76 12.9 Toxin HipA OS=Stenotrophomonas maltophilia OX=40324 GN=CEE60_13960 PE=4 SV=1
A0A245ZNB6 575 76 12.9 Toxin RTX-I translocation ATP-binding protein OS=Sphingomonas dokdonensis OX=344880 GN=apxIB PE=4 SV=1
A0A243MZU7 525 76 12.9 Toxin OS=Bacillus thuringiensis subsp. medellin OX=79672 GN=BK784_33475 PE=4 SV=1
A0A242M5O3 466 76 12.9 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM OS=Caballeronia sordidicola OX=196367 G
A0A239J188 938 121 12.9 Ca2+-binding protein, RTX toxin-related OS=Azospirillum sp. RU38E OX=1907313 GN=SAMN05880556_11820 PE=4
A0A239HLY5 729 94 12.9 Ca2+-binding protein, RTX toxin-related OS=Tropicimonas sediminicola OX=1031541 GN=SAMN05421757_103483 PE
A0A223MF23 464 76 12.9 Toxin and drug export protein A OS=Actinobacillus pleuropneumoniae OX=715 GN=tdeA PE=4 SV=1
A0A222U694 621 80 12.9 Toxin OS=Rhizobium leguminosarum bv. viciae OX=387 GN=CHY08_28250 PE=4 SV=1
A0A212AKH0 413 76 12.9 Toxin HipA OS=Haematobacter missouriensis OX=366616 GN=CDV52_16420 PE=4 SV=1
A0A209ARG9 441 76 12.9 Toxin HipA OS=Yersinia frederiksenii OX=29484 GN=CBW58_00860 PE=4 SV=1
A0A1Z3ZOA5 444 76 12.9 Toxin HipA OS=Burkholderia cepacia OX=292 GN=CEQ23_31635 PE=4 SV=1
A0A1Y1J2P3 457 76 12.9 Multidrug and toxin extrusion OS=Comamonas testosteroni OX=285 GN=CTR2_1550 PE=4 SV=1
A0A1YOFW73 448 76 12.9 Toxin HipA OS=Cellvibrio sp. PSBB006 OX=1987723 GN=CBR65_07430 PE=4 SV=1
A0A1X7CGS0 658 85 12.9 Predicted ribonuclease, toxin component of the YeeF-YezG toxin-antitoxin module OS=Bacillus sp. CC120222-
A0A1X6ZZ55 745 96 12.9 Toxin RTX-I translocation ATP-binding protein OS=Roseovarius gaetbuli OX=1356575 GN=apxIB_1 PE=4 SV=1
A0A1V6K507 576 76 12.9 Toxin RTX-I translocation ATP-binding protein OS=candidate division Hyd24-12 bacterium ADurb.Bin004 OX=18
A0A1V4K2B9 628 81 12.9 Multidrug and toxin extrusion protein OS=Patagioenas fasciata monilis OX=372326 GN=AV530_003348 PE=3 SV=1
A0A1T4ZJG4 667 86 12.9 Putative haemolysin-type calcium-binding toxin, RTX-like (Expressed) OS=Planktothrix sp. PCC 11201 OX=172
A0A1S8GXE0 396 76 12.9 Toxin OS=Pseudomonas sp. FSL W5-0299 OX=1917484 GN=BOO94_24075 PE=4 SV=1
A0A1S1VYP3 566 76 12.9 MFS toxin efflux pump OS=Colletotrichum incanum OX=1573173 GN=CSPAE12_03963 PE=4 SV=1
A0A1S1VTF8 735 95 12.9 Protoplast regeneration and killer toxin resistance protein OS=Colletotrichum incanum OX=1573173 GN=CSPAE
A0A1R4F9B6 477 76 12.9 Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM, homolog OS=Agrococcus casei LMG 22410
A0A1Q8RJ85 562 76 12.9 Putative HC-toxin efflux carrier TOXA 17 OS=Colletotrichum chlorophyti OX=708187 GN=CCHL11_05905 PE=4 SV=
A0A1Q3M733 421 76 12.9 Toxin HipA OS=Bacteroidales bacterium 45-6 OX=1895719 GN=BGN96_15045 PE=4 SV=1
A0A1N7K5Z8 735 95 12.9 AAA domain-containing protein, putative AbiEii toxin, Type IV TA system OS=Roseivivax lentus OX=633194 GN

C Scripts for allergen analysis

Script for making the identity over a window

```
from operator import itemgetter

def find_window(target_file):
    data_temp = open("results/temp.txt")
    index_list = [3, 4, 7, 8, 9, 10]
    for line in data_temp:
        if ">>>" in line:
            name = line[3:].rstrip('\n')
            continue
        if line.startswith(">>"):
            if 'pir' or 'prf' not in temp_list:
                temp_list = line.split("|")
                temp_list.pop()
                temp_list = "␣".join(temp_list)
                if 'pir' or 'prf' in temp_list:
                    temp_list = line.split()[0]
        if line.startswith("Smith-Waterman"):
            id_line = line.split()
            id_line = "␣".join(itemgetter(*index_list)(id_line))
            target_file.write(name + "\t" + temp_list[2:] + "\t" + id_line + "\n")
    target_file.close()
    data_temp.close()


def find_protein_and_identity(input_file, target_file):
    target = open('results/temp.txt', 'w')
    for line in input_file:
        if ">>>" in line:
            target.write(line[3:])
        if line.startswith(">>") or line.startswith("Smith-Waterman"):
            target.write(line)
    target.close()
    find_window(target_file)
```

% The script is invoked by executing a corresponding block in a iPython notebook, which calls the script with certain files.Afterwards, the resulting file is through a filter regarding if the data is scaled or not.

Listing 1: Filter function for 35% id over 80 amino acids

```
def filter_80aa_35id(input_file, target_file):
    write_counter = 0
    for line in input_file:
        # This if is a hack.. Better look back to this
        if 'ung' in line:
            temp = line.replace("%", "").split()
            identity = float(temp[5])
            overlap = float(temp[9])
            if identity >= 35.000 and overlap >= 80.0:
                target_file.write(line.replace("ungapped)", "") + "overlap")
                write_counter += 1
        else:
            temp = line.replace("%", "").split()
            identity = float(temp[5])
            overlap = float(temp[8])
            if identity >= 35.000 and overlap >= 80.0:
                target_file.write(line)
                write_counter += 1
    input_file.close()
    target_file.close()
    check_hits_with_counter(target_file.name, write_counter)
```

Listing 2: Filter function for 35% id over 80 amino acids (scaled)

```
def filter_80aa_35id_scaled(input_file, target_file):
    write_counter = 0
    for line in input_file:
        # This if is a hack.. Better look back to this
        if 'ung' in line:
            temp = line.replace("%", "").split()
            identity = float(temp[5])
            overlap = float(temp[9])
            if identity >= 35.000 and (identity*overlap)/80 >= 35.0 and overlap < 80:
                target_file.write(line.replace("ungapped)", "") + "overlap")
                write_counter += 1
        else:
            temp = line.replace("%", "").split()
            identity = float(temp[5])
            overlap = float(temp[8])
        if identity >= 35.0 and (identity*overlap)/80 >= 35.0 and overlap < 80:
            target_file.write(line)
            write_counter += 1
    input_file.close()
    target_file.close()
    check_hits_with_counter(target_file.name, write_counter)
```

Script for making Needleman-Wunsch alignment and comparison

For this part, we generate a bash script containing the sequences, that needs a full alignment test. The shell script is generated by the following function:

```
def make_sh_file(data_file, target_file, enzym, window, db):

    for line in data_file:
        allergen = line.replace("\n", "")
        fastagrep = '\n/z/linux/bin/fastagrep -t -p "%s" databases/%s.fasta > results/needle/%s.fasta' % (allergen,
            db, allergen)

        target_file.write(fastagrep)

        needle = "\nneedle -asequence %s.fasta -bsequence results/needle/%s.fasta -gapopen 10.0 -gapextend 0.5 -
            outfile results/needle/%s_%s.needle" % (enzym, allergen, enzym, allergen)

        target_file.write(needle)

    target_file.close()

    with open(target_file.name) as checker:
        counter = 0
        for line in checker:
            if enzym in line:
                counter += 1

    check_hits_with_counter(target_file.name, counter)
```

After that, the bash script is run via a system call.

D List of allergens from allergenonline

List of allergens that have been tested by the EFSA scientific opinion recommended allergen analysis described in section 2. The sequences were downloaded via `http://allergenonline.org`.

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Acarus siro	Mite	Aca s 13	Aero Mite	Acarus Aca s 13	131	ABL09307.1	118638268	9
Acarus siro	Mite	Unassigned	Aero Insect	Acarus siro Group 4 allergen	517	ABL09312.1	118638278	9
Actinidia arguta	Hardy Kiwi	Unassigned	Food Plant	Actinidia arguta kiwellin	213	AGC39172.1	441482362	14
Actinidia arguta	Hardy Kiwi	Unassigned	Food Plant	Actinidia arguta kiwellin	213	AGC39173.1	441482364	14
Actinidia arguta	Hardy Kiwi	Unassigned	Food Plant	Actinidia arguta kiwellin	213	AGC39174.1	441482366	14
Actinidia chinensis	Kiwi	Unassigned	Food Plant	Actinidia Act c 1 Act d 1 Actinidin	380	P00785.4	190358935	9
Actinidia chinensis	Kiwi	Act c 1	Food Plant	Actinidia Act c 10 LTP	15	P85204.1	378548410	13
Actinidia chinensis	Kiwi	Act c 5.0102	Food Plant	Actinidia Act c 5 kiwellin	213	AGC39168.1	441482354	14
Actinidia chinensis	Kiwi	Act c 8.0101	Food Plant	Actinidia Act c 8 Act d 8 PR-10	159	CAM31908.1	281552896	11
Actinidia chinensis	Kiwi	Unassigned	Unassigned	Actinidia Act d 12	462	ABB77213.1	82469930	16
Actinidia chinensis	Kiwi	Unassigned	Food Plant	Actinidia Act d 2 thaumatin like protein	20	P83958.1	68064399	7
Actinidia chinensis	Kiwi	Unassigned	Food Plant	Actinidia Act d 2 thaumatin like protein	225	AGC39176.1	441482370	14
Actinidia deliciosa	Kiwi	Act d 1	Food Plant	Actinidia Act c 1 Act d 1 Actinidin	380	CAA34486.1	15984	7
Actinidia deliciosa	Kiwi	Unassigned	Food Plant	Actinidia Act c 1 Act d 1 Actinidin	380	AAA32629.1	166317	7
Actinidia deliciosa	Kiwi	Unassigned	Food Plant	Actinidia Act c 1 Act d 1 Actinidin	380	A5HII1.1	193806686	12
Actinidia deliciosa	Kiwi	Act d 8.0101	Food Plant	Actinidia Act c 8 Act d 8 PR-10	157	CAM31909.1	281552898	11
Actinidia deliciosa	Kiwi	Act d 10.0201	Food Plant	Actinidia Act d 10 LTP	92	P85206.1	378548411	13
Actinidia deliciosa	Kiwi	Act d 10.0101	Food Plant	Actinidia Act d 10 LTP	92	P86137.2	378405189	13
Actinidia deliciosa	Kiwi	Act d 11	Food Plant	Actinidia Act d 11 Kirola MLP	150	P85524.1	332319679	12

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Actinidia deliciosa	Kiwi	Act d 2.0101	Food Plant	Actinidia Act d 2 thaumatin like protein	225	CAI38795.2	71057064	7
Actinidia deliciosa	Kiwi	Unassigned	Food Plant	Actinidia Act d 2 thaumatin like protein	201	ABQ42566.1	146737976	9
Actinidia deliciosa	Kiwi	Act d 4.0101	Food Plant	Actinidia Act d 4 Phytocystatin	116	AAR92223.1	40807635	7
Actinidia deliciosa	Kiwi	Act d 5.0101	Food Plant	Actinidia Act d 5 kiwellin	189	P84527.1	85701136	7
Actinidia deliciosa	Kiwi	Unassigned	Food Plant	Actinidia Act d 5 kiwellin	213	AGC39164.1	441482346	14
Actinidia deliciosa	Kiwi	Unassigned	Food Plant	Actinidia Act d 5 kiwellin	213	AGC39165.1	441482348	14
Actinidia deliciosa	Kiwi	Unassigned	Food Plant	Actinidia Act d 5 kiwellin	213	AGC39166.1	441482350	14
Actinidia deliciosa	Kiwi	Unassigned	Food Plant	Actinidia Act d 5 kiwellin	213	AGC39167.1	441482352	14
Actinidia deliciosa	Kiwi	Unassigned	Food Plant	Actinidia Act d 5 kiwellin	189	4X9U_B	906848988	17
Actinidia deliciosa	Kiwi	Act d 9.0101	Food Plant	Actinidia Act d 9, profilin	109		100	16
Actinidia eriantha	Climber (plant)	Unassigned	Food Plant	Actinidia eriantha kiwellin	213	AGC39169.1	441482356	14
Actinidia eriantha	Climber (plant)	Unassigned	Food Plant	Actinidia eriantha kiwellin	213	AGC39170.1	441482358	14
Actinidia eriantha	Climber (plant)	Unassigned	Food Plant	Actinidia eriantha kiwellin	213	AGC39171.1	441482360	14
Aedes aegypti	Yellow fever mosquito	Aed a 1	Venom or Salivary	Aedes Aed a 1 apyrase	562	AAC37218.1	556272	7
Aedes aegypti	Yellow fever mosquito	Unassigned	Venom or Salivary	Aedes Aed a 1 apyrase	562	P50635.2	193806340	10
Aedes aegypti	Yellow fever mosquito	Aed a 11	Unassigned	Aedes Aed a 11 Lysosomal protease	387	XP_001657556.2	1218215869	18
Aedes aegypti	Yellow fever mosquito	Aed a 2	Venom or Salivary	Aedes Aed a 2	321	P18153.2	205525919	9
Aedes aegypti	Yellow fever mosquito	Aed a 3	Venom or Salivary	Aedes Aed a 3	253	AAB58417.1	2114497	7
Aedes aegypti	Yellow fever mosquito	Unassigned	Venom or Salivary	Aedes Aed a 3	273	ABF18122.1	94468546	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Aedes aegypti	Yellow fever mosquito	Aed a 5.0101	Unassigned	Aedes Aed a 5 Sarcoplasmic Ca+ bind	191	XP_001653462.1	157119961	17
Aedes aegypti	Yellow fever mosquito	Aed a 6.0101	Unassigned	Aedes Aed a 6 Porin 3	282	XP_001654143.1	157124666	17
Aedes aegypti	Yellow fever mosquito	Aed a 7.0101	Unassigned	Aedes Aed a 7	204	XP_001654291.1	157125324	17
Aedes aegypti	Yellow fever mosquito	Aed a 8.0101	Unassigned	Aedes Aed a 8 HSP70	655	ABF18258.1	94468818	17
Aedes aegypti	Yellow fever mosquito	Aed a 10.0201	Unassigned	Aedes aegypti Aed a 10	284	XP_001655948.1	157131813	16
Aedes aegypti	Yellow fever mosquito	Aed a 10.0101	Unassigned	Aedes aegypti Aed a 10	285	XP_001655954.1	157131825	16
Aedes aegypti	Yellow fever mosquito	Unassigned	Unassigned	Aedes aegypti Aed a 4 alpha glucosidase	579	P13080.1	126713	17
Agrostis alba	Bent grass	Unassigned	Aero Plant	Agrostis Agr a 1	26	E37396	320606	7
Agrostis alba	Bent grass	Unassigned	Aero Plant	Agrostis Agr a 1	35	Q7M1X7	75139987	7
Agrostis alba	Bent grass	Unassigned	Aero Plant	Agrostis Agr a 1	35	Q7M1X9	75139989	7
Alnus glutinosa	Alder	Aln g 1	Aero Plant	Alnus Aln g 1	160	AAB24432.1	261407	7
Alnus glutinosa	Alder	Aln g 1.0101	Aero Plant	Alnus Aln g 4	85	CAA76831.1	3319651	7
Alternaria alternata	Fungus	Alt a 1.0101	Aero Fungi	Alternaria Alt a 1	157	AAB47552.1	1842045	7
Alternaria alternata	Fungus	Unassigned	Aero Fungi	Alternaria Alt a 1	115	AAM77471.1	21913174	7
Alternaria alternata	Fungus	Alt a 1.0102	Aero Fungi	Alternaria Alt a 1	157	AAS75297.1	45680856	7
Alternaria alternata	Fungus	Unassigned	Aero Fungi	Alternaria Alt a 1	133	3V0R_A	390980892	13
Alternaria alternata	Fungus	Unassigned	Aero Fungi	Alternaria Alt a 1	130	4AUD_B	508123617	15
Alternaria alternata	Fungus	Alt a 10.0101	Aero Fungi	Alternaria Alt a 10 ADH	497	CAA55071.2	76666767	7
Alternaria alternata	Fungus	Alt a 12	Aero Fungi	Alternaria Alt a 12 Ribosomal BP P1	110	P49148.1	1350779	7
Alternaria alternata	Fungus	Alt a 13.0101	Aero Fungi	Alternaria Alt a 13	231	Q6R4B4.1	74611808	10

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Alternaria alternata	Fungus	Alt a 3	Aero Fungi	Alternaria Alt a 3 HSP	152	P78983.2	14423730	7
Alternaria alternata	Fungus	Alt a 4	Aero Fungi	Alternaria Alt a 4 thioredoxin	436	Q00002.2	85701160	7
Alternaria alternata	Fungus	Alt a 5	Aero Fungi	Alternaria Alt a 5 ribosomal P2	113	AAB48041.1	1850540	7
Alternaria alternata	Fungus	Unassigned	Aero Fungi	Alternaria Alt a 5 ribosomal P2	113	P42037.1	1173071	10
Alternaria alternata	Fungus	Alt a 6	Aero Fungi	Alternaria Alt a 6 enolase	438	Q9HDT3.2	14423684	7
Alternaria alternata	Fungus	Alt a 7.0101	Aero Fungi	Alternaria Alt a 7 flavodoxin	204	P42058.1	1168402	9
Alternaria alternata	Fungus	Unassigned	Unassigned	Alternaria Alt a 7 flavodoxin	261	OWY50380.1	1213711549	18
Alternaria alternata	Fungus	Alt a 8.0101	Aero Fungi	Alternaria Alt a 8 (mannitol dehydrogenase)	266	AAO91800.1	37780013	8
Alternaria alternata	Fungus	Unassigned	Aero Fungi	Alternaria Alt a 8 (mannitol dehydrogenase)	266	P0C0Y4.2	118595439	8
Alternaria alternata	Fungus	Alt a 14.0101	Aero Fungi	Alternaria MnSOD Alt a 14	191	AGS80276.1	529279957	15
Alternaria alternata	Fungus	Unassigned	Aero Fungi	Alternaria Nuc Transport 2	124	CAD38167.1	21748153	7
Alternaria alternata	Fungus	Unassigned	Aero Fungi	Alternaria TCTP IgE binding	169	ABI26088.1	112824341	11
Amaranthus retroflexus	Common Amaranth	Ama r 2.0101	Aero Plant	Amaranthus Ama r 2 Proflin	133	ACP43298.1	227937304	10
Amaranthus retroflexus	Common Amaranth	Unassigned	Unassigned	Amaranthus retroflexus Ama r 1	168	AKV72168.1	914410010	16
Ambrosia artemisiifolia	Short ragweed	Amb a 1.0101	Aero Plant	Ambrosia Amb a 1	396	P27759.1	113475	7
Ambrosia artemisiifolia	Short ragweed	Amb a 1.0201	Aero Plant	Ambrosia Amb a 1	398	P27760.1	113476	7
Ambrosia artemisiifolia	Short ragweed	Amb a 1.0301	Aero Plant	Ambrosia Amb a 1	397	P27761.1	113477	7
Ambrosia artemisiifolia	Short ragweed	Amb a 1.0401	Aero Plant	Ambrosia Amb a 1	392	P28744.1	113478	7
Ambrosia artemisiifolia	Short ragweed	Amb a 1.0303	Aero Plant	Ambrosia Amb a 1	397	AAA32669.1	166443	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 1	396	CBW30986.1	302127810	12
Ambrosia artemisiifolia	Short ragweed	Amb a 1.0202	Aero Plant	Ambrosia Amb a 1	398	CBW30987.1	302127812	12
Ambrosia artemisiifolia	Short ragweed	Amb a 1.0304	Aero Plant	Ambrosia Amb a 1	397	CBW30988.1	302127814	12
Ambrosia artemisiifolia	Short ragweed	Amb a 1.0305	Aero Plant	Ambrosia Amb a 1	397	CBW30989.1	302127816	12
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 1	397	CBW30990.1	302127818	12
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 1	397	CBW30991.1	302127820	12
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 1	397	CBW30992.1	302127822	12
Ambrosia artemisiifolia	Short ragweed	Amb a 1.0402	Aero Plant	Ambrosia Amb a 1	387	CBW30993.1	302127824	12
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 1	397	CBW30994.1	302127826	12
Ambrosia artemisiifolia	Short ragweed	Amb a 1.0502	Aero Plant	Ambrosia Amb a 1	397	CBW30995.1	302127828	12
Ambrosia artemisiifolia	Short ragweed	Amb a 10.0101	Aero Plant	Ambrosia Amb a 10	160	AAX77686.1	62249491	7
Ambrosia artemisiifolia	Short ragweed	Amb a 2	Aero Plant	Ambrosia Amb a 2	397	P27762.1	113479	7
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 4	164	CBJ24286.1	285005079	11
Ambrosia artemisiifolia	Short ragweed	Amb a 4.0101	Aero Plant	Ambrosia Amb a 4	164	CBK52317.1	291197394	12
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 4	111	CBK62693.1	291482306	12
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 4	140	CBK62694.1	291482308	12
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 4	134	CBK62695.1	291482310	12
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 4	96	CBK62697.1	291482314	12
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 4	110	CBK62698.1	291482316	12

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 4	116	CBK62699.1	291482318	12
Ambrosia artemisiifolia	Short ragweed	Amb a 6	Aero Plant	Ambrosia Amb a 6	118	O04004.1	14285595	7
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 8 profilin	133	AAP15203.1	34851182	7
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 8 profilin	131	AAP15202.1	34851180	7
Ambrosia artemisiifolia	Short ragweed	Unassigned	Aero Plant	Ambrosia Amb a 8 profilin	131	AAP15201.1	34851178	7
Ambrosia artemisiifolia	Short ragweed	Amb a 8.0101	Aero Plant	Ambrosia Amb a 8 profilin	133	AAX77687.1	62249502	7
Ambrosia artemisiifolia	Short ragweed	Amb a 8.0102	Aero Plant	Ambrosia Amb a 8 profilin	133	AAX77688.1	62249512	7
Ambrosia artemisiifolia	Short ragweed	Unassigned	Unassigned	Ambrosia Amb a 8 profilin	135	5EM1_A	1035439203	18
Ambrosia artemisiifolia	Short ragweed	Unassigned	Unassigned	Ambrosia Amb a 8 profilin	134	5EV0_B	1035439209	18
Ambrosia artemisiifolia	Short ragweed	Amb a 9.0101	Aero Plant	Ambrosia Amb a 9	83	AAX77684.1	62249470	7
Ambrosia artemisiifolia	Short ragweed	Amb a 9.0102	Aero Plant	Ambrosia Amb a 9	83	AAX77685.1	62249481	7
Ambrosia artemisiifolia	Short ragweed	Amb a 11.0101	Unassigned	Ambrosia artemisiifolia Amb a 11	386	AHA56102.1	558482540	15
Ambrosia artemisiifolia	Short ragweed	Unassigned	Unassigned	Ambrosia artemisiifolia Amb a 11	385	5EGW_B	1023176264	17
Ambrosia artemisiifolia (elatior)	Short ragweed	Amb a 3	Aero Plant	Ambrosia Amb a 3	101	P00304.2	416636	7
Ambrosia artemisiifolia (elatior)	Short ragweed	Amb a 5	Aero Plant	Ambrosia Amb a 5 Ra5	45	P02878.1	114090	7
Ambrosia psilostachya	Western ragweed	Amb p 5.0101	Aero Plant	Ambrosia Amb a 5 Ra5	77	AAA20065.1	515953	7
Ambrosia psilostachya	Western ragweed	Unassigned	Aero Plant	Ambrosia Amb a 5 Ra5	77	AAA20067.1	515954	7
Ambrosia psilostachya	Western ragweed	Amb p 5.0201	Aero Plant	Ambrosia Amb a 5 Ra5	77	AAA20064.1	515955	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Ambrosia psilostachya	Western ragweed	Unassigned	Aero Plant	Ambrosia Amb a 5 Ra5	77	AAA20066.1	515956	7
Ambrosia psilostachya	Western ragweed	Unassigned	Aero Plant	Ambrosia Amb a 5 Ra5	77	AAA20068.1	515957	7
Ambrosia trifida	Giant ragweed	Amb t 5	Aero Plant	Ambrosia Amb t 5 Ra5G	73	P10414.2	114091	7
Amphioctopus fangsiao	Octopus	Unassigned	Food Animal	Amphioctopus arginine kinase	348	AEK65120.1	340742817	12
Anacardium occidentale	Cashew	Ana 0 1.0102	Food Plant	Anacardium Ana o 1	536	AAM73729.1	21666498	7
Anacardium occidentale	Cashew	Ana 0 1.0101	Food Plant	Anacardium Ana o 1	538	AAM73730.2	21914823	7
Anacardium occidentale	Cashew	Ana o 2	Food Plant	Anacardium Ana o 2	457	AAN76862.1	25991543	7
Anacardium occidentale	Cashew	Ana o 3	Food Plant	Anacardium Ana o 3	138	AAL91665.1	24473800	7
Ananas comosus	Pineapple	Ana c 2.0101	Aero Plant	Ananas Ana c 2 Bromelain precursor	351	O23791.1	75277440	7
Ananas comosus	Pineapple	Ana c 1.0101	Food Plant	Ananas profilin Ana c 1	131	Q94JN2.1	75306610	10
Anaphe panda		Unassigned	Unassigned	Thaumetopoea Tha p 2	84	CDZ09832.1	925059182	17
Anisakis pegreffii	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 12	264	AGC60026.1	442577845	14
Anisakis pegreffii	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 12	264	AGC60027.1	442577847	14
Anisakis pegreffii	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 12	264	AGC60028.1	442577849	14
Anisakis pegreffii	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 2 paramyosin	869	AGC60020.1	442577833	14
Anisakis simplex	Parasitic fish worm	Ani s 1	Food Animal	Anisakis Ani s 1 protease inhibitor	194	Q7Z1K3.1	47605452	7
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 1 protease inhibitor	163	AGC60035.1	442577863	14
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 1 protease inhibitor	163	AGC60036.1	442577865	14
Anisakis simplex	Parasitic fish worm	Ani s 10.0101	Food Animal	Anisakis Ani s 10	231	ACZ95445.1	272574378	11

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Anisakis simplex	Parasitic fish worm	Ani s 11.0101	Food Animal	Anisakis Ani s 11	307	BAJ78220.1	323575361	12
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 11	160	BAJ78221.1	323575363	12
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 11	287	BAJ78222.1	323575365	12
Anisakis simplex	Parasitic fish worm	Ani s 12.0101	Food Animal	Anisakis Ani s 12	295	BAJ78223.1	323575367	12
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 12	264	AGC60029.1	442577851	14
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 12	264	AGC60030.1	442577853	14
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 12	264	AGC60031.1	442577855	14
Anisakis simplex	Parasitic fish worm	Unassigned	Unassigned	Anisakis Ani s 14	217	BAT62430.1	957554293	17
Anisakis simplex	Parasitic fish worm	Ani s 2	Food Animal	Anisakis Ani s 2 paramyosin	473	AAF75225.1	8453086	7
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 2 paramyosin	869	Q9NJA9.1	42559536	9
Anisakis simplex	Parasitic fish worm	Ani s 3	Food Animal	Anisakis Ani s 3 tropomyosin	284	Q9NAS5.1	14423976	7
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 3 tropomyosin	284	AEQ28167.1	350285785	13
Anisakis simplex	Parasitic fish worm	Ani s 4	Food Animal	Anisakis Ani s 4	14	P83885.1	47605398	7
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 4	115	CAK50389.1	110346534	8
Anisakis simplex	Parasitic fish worm	Ani s 5.0101	Food Animal	Anisakis Ani s 5 SXP/RAL-2 family protein	152	BAF43534.1	121308878	8
Anisakis simplex	Parasitic fish worm	Ani s 7.0101	Food Animal	Anisakis Ani s 7 UA3-recognized allergen	1096	ABL77410.1	119524036	9
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 8 SXP/RAL-2 family protein 2	150	BAF75681.1	155676636	9
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 8 SXP/RAL-2 family protein 2	150	BAF75704.1	155676682	9
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 8 SXP/RAL-2 family protein 2	150	BAF75705.1	155676684	9

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 8 SXP/RAL-2 family protein 2	150	BAF75706.1	155676686	9
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 8 SXP/RAL-2 family protein 2	150	BAF75707.1	155676688	9
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 8 SXP/RAL-2 family protein 2	150	BAF75708.1	155676690	9
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 8 SXP/RAL-2 family protein 2	150	BAF75709.1	155676692	9
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 8 SXP/RAL-2 family protein 2	150	BAF75710.1	155676694	9
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 8 SXP/RAL-2 family protein 2	150	BAF75711.1	155676696	9
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis Ani s 8 SXP/RAL-2 family protein 2	150	BAF75712.1	155676698	9
Anisakis simplex	Parasitic fish worm	Ani s 9.0101	Food Animal	Anisakis Ani s 9	147	ABV55106.1	157418806	9
Anisakis simplex	Parasitic fish worm	Unassigned	Food Animal	Anisakis simplex troponin-like	161	CAB58171.1	6065738	7
Anthoxanthum odoratum	Sweet vernal grass	Unassigned	Aero Plant	Anthoxanthum Ant o 1	26	G37396	320607	7
Anthoxanthum odoratum	Sweet vernal grass	Ant o 1.0101	Aero Plant	Anthoxanthum Ant o 1	32	Q7M1X6	75139986	7
Anthoxanthum odoratum	Sweet vernal grass	Unassigned	Aero Plant	Anthoxanthum Ant o 1	32	Q7M1Y0	75139990	7
Apis cerana	Indian honeybee	Unassigned	Venom or Salivary	Apis Api m 1 Api d 1 Api c 1	134	A59055	7435005	7
Apis cerana cerana	Indian honeybee	Api c 1.0101	Venom or Salivary	Apis Api m 1 Api d 1 Api c 1	134	AAK09361.1	12958583	15
Apis dorsata	Giant honeybee	Api d 1.0101	Venom or Salivary	Apis Api m 1 Api d 1 Api c 1	134	Q7M4I5.1	47117012	7
Apis dorsata	Giant honeybee	Unassigned	Venom or Salivary	Apis Api m 4 Melittin	26	P01502.1	126955	7
Apis mellifera	Honeybee	Api m 1	Venom or Salivary	Apis Api m 1 Api d 1 Api c 1	167	P00630.3	24418862	7
Apis mellifera	Honeybee	Unassigned	Venom or Salivary	Apis Api m 10 icarapin	223	ABF21077.1	94471622	7
Apis mellifera	Honeybee	Api m 10.0101	Venom or Salivary	Apis Api m 10 icarapin	175	ABF21078.1	94471624	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Apis mellifera	Honeybee	Api m 2	Venom or Salivary	Apis Api m 2	382	Q08169.1	585279	7
Apis mellifera	Honeybee	Unassigned	Venom or Salivary	Apis Api m 3 acid phosphatase	388	ACI25605.1	208342441	10
Apis mellifera	Honeybee	Api m 3.0101	Venom or Salivary	Apis Api m 3 acid phosphatase	388	Q5BLY5.1	74835477	12
Apis mellifera	Honeybee	Api m 4.0101	Venom or Salivary	Apis Api m 4 Melittin	70	CAA26038.1	5622	7
Apis mellifera	Honeybee	Unassigned	Venom or Salivary	Apis Api m 4 Melittin	27	MEHB2	69552	7
Apis mellifera	Honeybee	Api m 5.0101	Venom or Salivary	Apis Api m 5 dipeptidylpeptidase	775	NP_001119715.1	187281543	15
Apis mellifera	Honeybee	Unassigned	Venom or Salivary	Apis Api m 6	92	NP_001035360.1	94400907	7
Apis mellifera	Honeybee	Unassigned	Venom or Salivary	Apis Api m 6	94	ABD51779.1	88770352	10
Apis mellifera	Honeybee	Api m 11.0101	Venom or Salivary	Apis mellifera Api m 11	416	NP_001011564.1	58585070	15
Apis mellifera	Honeybee	Api m 11.0201	Venom or Salivary	Apis mellifera Api m 11	423	AAY21180.1	62910925	15
Apis mellifera	Honeybee	Api m 12.0101	Venom or Salivary	Apis mellifera Api m 12	1770	CAD56944.1	29329817	15
Apis mellifera carnica	Honeybee	Unassigned	Venom or Salivary	Apis Api m 10 icarapin	12	AHM25038.1	594708629	16
Apis mellifera carnica	Honeybee	Unassigned	Venom or Salivary	Apis Api m 10 icarapin	19	AHM25037.1	594708627	16
Apis mellifera carnica	Honeybee	Unassigned	Venom or Salivary	Apis Api m 10 icarapin	25	AHM25036.1	594708625	16
Apis mellifera carnica	Honeybee	Unassigned	Venom or Salivary	Apis Api m 10 icarapin	41	AHM25035.1	594708623	16
Apium graveolens	Celery	Api g 1.0101	Food Plant	Apium Api g 1	154	P49372.1	1346568	7
Apium graveolens	Celery	Api g 1.0201	Food Plant	Apium Api g 1	159	P92918.1	14423646	9
Apium graveolens	Celery	Api g 2.0101	Food Plant	Apium Api g 2	118	ACV04796.1	256600126	12
Apium graveolens	Celery	Api g 4	Food Plant	Apium Api g 4	134	AAD29409.1	4761578	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Apium graveolens	Celery	Api g 5.0101	Food Plant	Apium Api g 5	86	P81943.3	33300920	10
Apium graveolens Rapaceum Group	Celery	Api g 6.0101	Food Plant	Apium graveolens Api g 6 LTP 2	67	P86809.1	550540827	15
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Agglutinin (lectin)	273	AAB22817.1	253289	7
Arachis hypogaea	Peanut	Ara h 1	Food Plant	Arachis Ara h 1	614	P43237.1	1168390	7
Arachis hypogaea	Peanut	Ara h 1	Food Plant	Arachis Ara h 1	626	P43238.1	1168391	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 1	299	AAT00595.1	46560474	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 1	303	AAT00594.1	46560472	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 1	428	AAT00596.1	46560476	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 1	619	ADQ53858.1	312233063	12
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 1	418	3SMH_A	375332427	13
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 1	418	3S7E_A	347447588	13
Arachis hypogaea	Peanut	Ara h 12.0101	Unassigned	Arachis Ara h 12	71	B3EWP3.1	1018736824	17
Arachis hypogaea	Peanut	Unassigned	Unassigned	Arachis Ara h 13 defensin	72	C0HJZ1.1	1018736837	17
Arachis hypogaea	Peanut	Unassigned	Unassigned	Arachis Ara h 13 defensin	79	B3EWP4.1	1018736830	17
Arachis hypogaea	Peanut	Ara h 2.0201	Food Plant	Arachis Ara h 2	172	AAN77576.1	26245447	7
Arachis hypogaea	Peanut	Ara h 2.0101	Food Plant	Arachis Ara h 2	169	AAM78596.1	31322017	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 2	156	AAK96887.1	15418705	10
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 2	158	ACN62248.1	224747150	10

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Arachis hypogaea	Peanut	Ara h 3.0101	Food Plant	Arachis Ara h 3 Glycinin	507	AAC63045.1	3703107	7
Arachis hypogaea	Peanut	Ara h 3.0201	Food Plant	Arachis Ara h 3 Glycinin	530	AAD47382.1	5712199	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 3 Glycinin	538	AAM46958.1	21314465	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 3 Glycinin	219	AAM93157.1	22135348	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 3 Glycinin	512	ABI17154.1	112380623	8
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 3 Glycinin	530	ACH91862.1	199732457	10
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 3 Glycinin	510	3C3V_A	224036293	10
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 3 Glycinin	512	ADQ53859.1	312233065	12
Arachis hypogaea	Peanut	Ara h 5	Food Plant	Arachis Ara h 5	131	AAD55587.1	5902968	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 5	131	ADB96066.1	284810529	11
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 5	131	AGA84056.1	431812555	14
Arachis hypogaea	Peanut	Ara h 6	Food Plant	Arachis Ara h 6	129	AAD56337.1	5923742	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 6	144	AAL37561.1	17225991	7
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 6	127	1W2Q_A	159163254	9
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 6	145	Q647G9.1	75114094	10
Arachis hypogaea	Peanut	Ara h 7.0101	Food Plant	Arachis Ara h 7	160	AAD56719.1	5931948	7
Arachis hypogaea	Peanut	Ara h 7.0201	Food Plant	Arachis Ara h 7	164	ABW17159.1	158121995	10
Arachis hypogaea	Peanut	Ara h 8.0101	Food Plant	Arachis Ara h 8	157	AAQ91847.1	37499626	7
Arachis hypogaea	Peanut	Ara h 8.0201	Food Plant	Arachis Ara h 8	153	ABP97433.1	145904610	9

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 8	157	ACA79908.1	169786740	9
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis Ara h 8	157	ABG85155.1	110676574	12
Arachis hypogaea	Peanut	Ara h 9.0101	Food Plant	Arachis Ara h 9 LTP isoallergens	116	ABX56711.1	161087230	10
Arachis hypogaea	Peanut	Ara h 9.0201	Food Plant	Arachis Ara h 9 LTP isoallergens	92	ABX75045.1	161610580	10
Arachis hypogaea	Peanut	Ara h 10.0101	Food Plant	Arachis hypogaea Ara h 10	169	AAU21499.2	113200509	15
Arachis hypogaea	Peanut	Ara h 10.0102	Food Plant	Arachis hypogaea Ara h 10	150	AAU21500.1	52001239	15
Arachis hypogaea	Peanut	Ara h 11.0101	Food Plant	Arachis hypogaea Ara h 11	137	AAZ20276.1	71040655	15
Arachis hypogaea	Peanut	Unassigned	Food Plant	Arachis hypogaea Ara h 11	137	Q45W86	122218540	16
Argas reflexus	European pigeon tick	Arg r 1	Venom or Salivary	Argas Arg r 1	159	CAG26895.1	58371884	7
Argas reflexus	European pigeon tick	Unassigned	Venom or Salivary	Argas Arg r 1	144	2X45_A	322812205	12
Artemisia absinthium		Unassigned	Unassigned	Artemisia Art v 1	108	AHF71021.1	573005946	17
Artemisia annua		Unassigned	Unassigned	Artemisia Art v 1	108	AHF71022.1	573005948	17
Artemisia californica		Unassigned	Unassigned	Artemisia Art v 1	108	AHF71023.1	573005950	17
Artemisia frigida		Unassigned	Unassigned	Artemisia Art v 1	108	AHF71024.1	573005952	17
Artemisia ludoviciana		Unassigned	Unassigned	Artemisia Art v 1	108	AHF71025.1	573005954	17
Artemisia tridentata		Unassigned	Unassigned	Artemisia Art v 1	108	AHF71026.1	573005956	17
Artemisia vulgaris	Mugwort	Art v 1	Aero Plant	Artemisia Art v 1	132	AAO24900.1	27818335	7
Artemisia vulgaris	Mugwort	Art v 2.0101	Aero Plant	Artemisia Art v 2	162	CAK50834.1	148887203	9
Artemisia vulgaris	Mugwort	Art v 3.0101	Aero Plant	Artemisia Art v 3	37	P0C088.1	73621307	7
Artemisia vulgaris	Mugwort	Art v 3.0201	Aero Plant	Artemisia Art v 3	114	ACE07186.1	189544578	11
Artemisia vulgaris	Mugwort	Art v 3.0202	Aero Plant	Artemisia Art v 3	116	ACE07187.1	189544584	11

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Artemisia vulgaris	Mugwort	Art v 3.0301	Aero Plant	Artemisia Art v 3	117	ACE07188.1	189544590	11
Artemisia vulgaris	Mugwort	Unassigned	Aero Plant	Artemisia Art v 3	117	ACE07189.1	189544595	11
Artemisia vulgaris	Mugwort	Art v 4.0101	Aero Plant	Artemisia Art v 4	133	CAD12861.1	25955969	15
Artemisia vulgaris	Mugwort	Art v 4.0201	Aero Plant	Artemisia Art v 4	133	CAD12862.1	25955971	15
Artemisia vulgaris	Mugwort	Unassigned	Unassigned	Artemisia Art v 4	135	5EM0_A	1035439202	18
Artemisia vulgaris	Mugwort	Art v 6.0101	Aero Plant	Artemisia Art v 6 pectate lyase	396	AAX85388.1	62530263	8
Artemisia vulgaris	Mugwort	Art v 5.0101	Aero Plant	Artemisia mugwort Art v 5	82	AAX85389.1	62530265	15
Arthroderma benhamiae	Fungus	Unassigned	Contact	Trichophyton (Arthroderma) Tri m 4	726	CAD23611.1	23894232	7
Arthroderma benhamiae	Fungus	Unassigned	Contact	Trichophyton (Arthroderma) Tri r 2	292	CAD23613.1	23894240	7
Arthroderma benhamiae	Fungus	Unassigned	Contact	Trichophyton (Arthroderma) Tri r 2	404	CAD23614.1	23894244	7
Arthroderma vanbreuseghemii	Fungus	Unassigned	Contact	Trichophyton (Arthroderma) Tri m 4	726	BAH09387.1	219687753	10
Ascaris lumbricoides	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	134	AAD13644.1	2735096	7
Ascaris lumbricoides	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	134	AAD13645.1	2735098	7
Ascaris lumbricoides	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	133	AAD13647.1	2735102	7
Ascaris lumbricoides	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	133	AAD13649.1	2735106	7
Ascaris lumbricoides	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	267	AAD13650.1	2735108	7
Ascaris lumbricoides	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	267	AAD13651.1	2735110	7
Ascaris lumbricoides	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	267	AAD13652.1	2735112	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Ascaris lumbricoides	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	134	AAB93837.1	2735114	7
Ascaris lumbricoides	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	134	AAB93839.1	2735118	7
Ascaris lumbricoides	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	134	AAD13646.1	2735100	7
Ascaris lumbricoides	Parasitic roundworm	Asc l 3.0101	Worm (parasite)	Ascaris tropomyosin Asc l 3	287	ACN32322.1	224016002	10
Ascaris suum	Parasitic roundworm	Asc s 1	Worm (parasite)	Ascaris Asc s 1	68	AAB26195.1	299550	7
Ascaris suum	Parasitic roundworm	Asc s 1	Worm (parasite)	Ascaris Asc s 1	1365	Q06811.2	77416849	7
Ascaris suum	Parasitic roundworm	Unassigned	Worm (parasite)	Ascaris Asc s 1	134	2XV9_A	343197079	12
Ascaris suum	Parasitic roundworm	Asc s 13.0101	Worm (parasite)	Ascaris lumbricoides/suum Glutathione S-transfera	206	P46436.3	1170109	15
Aspergillus flavus	Fungus	Unassigned	Aero Fungi	Aspergillus Oryzin Asp o 13, fl 13	403	Q9UVU3	74665726	7
Aspergillus fumigatus	Fungus	Asp f 1	Aero Fungi	Aspergillus Asp f 1	125	CAA06305.1	3021324	7
Aspergillus fumigatus	Fungus	Asp f 1	Aero Fungi	Aspergillus Asp f 1	150	AAF86369.1	9280360	7
Aspergillus fumigatus	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 1	176	P67875.1	54039254	7
Aspergillus fumigatus	Fungus	Asp f 10	Aero Fungi	Aspergillus Asp f 10	395	CAA59419.1	963013	7
Aspergillus fumigatus	Fungus	Asp f 11	Aero Fungi	Aspergillus Asp f 11	178	CAB44442.1	5019414	7
Aspergillus fumigatus	Fungus	Asp f 18.0101	Aero Fungi	Aspergillus Asp f 18 and Asp n 18	495	CAA73782.1	2143220	7
Aspergillus fumigatus	Fungus	Asp f 2	Aero Fungi	Aspergillus Asp f 2	250	AAB07620.1	664852	7
Aspergillus fumigatus	Fungus	Asp f 2	Aero Fungi	Aspergillus Asp f 2	310	P79017.2	83300352	7
Aspergillus fumigatus	Fungus	Asp f 22	Aero Fungi	Aspergillus Asp f 22	438	AAK49451.1	13925873	7
Aspergillus fumigatus	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 22	438	Q96X30.3	83288046	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Aspergillus fumigatus	Fungus	Asp f 23.0101	Aero Fungi	Aspergillus Asp f 23	392	AAM43909.1	21215170	7
Aspergillus fumigatus	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 23	392	Q8NKF4.2	83305621	7
Aspergillus fumigatus	Fungus	Asp f 27.0101	Aero Fungi	Aspergillus Asp f 27	163	CAI78448.1	91680605	7
Aspergillus fumigatus	Fungus	Asp f 28.0101	Aero Fungi	Aspergillus Asp f 28	108	CAI78449.1	91680607	7
Aspergillus fumigatus	Fungus	Asp f 29.0101	Aero Fungi	Aspergillus Asp f 29	110	CAI78450.1	91680609	7
Aspergillus fumigatus	Fungus	Asp f 3	Aero Fungi	Aspergillus Asp f 3	168	AAB95638.1	2769700	7
Aspergillus fumigatus	Fungus	Asp f 34.0101	Aero Fungi	Aspergillus Asp f 34	185	CAM54066.1	133920236	8
Aspergillus fumigatus	Fungus	Asp f 4	Aero Fungi	Aspergillus Asp f 4	286	CAA04959.1	3005839	7
Aspergillus fumigatus	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 4	322	O60024.2	83300369	7
Aspergillus fumigatus	Fungus	Asp f 5	Aero Fungi	Aspergillus Asp f 5	634	CAA83015.1	3776613	7
Aspergillus fumigatus	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 5	634	P46075.3	85541646	11
Aspergillus fumigatus	Fungus	Asp f 6	Aero Fungi	Aspergillus Asp f 6	221	AAB60779.1	1648970	7
Aspergillus fumigatus	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 6	210	Q92450.3	83305645	7
Aspergillus fumigatus	Fungus	Asp f 7	Aero Fungi	Aspergillus Asp f 7	270	O42799.2	83300389	7
Aspergillus fumigatus	Fungus	Asp f 8	Aero Fungi	Aspergillus Asp f 8	111	CAB64688.1	6686524	7
Aspergillus fumigatus	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 8	111	Q9UUZ6.2	83305635	7
Aspergillus fumigatus	Fungus	Asp f 9	Aero Fungi	Aspergillus Asp f 9	302	CAA11266.1	2879890	7
Aspergillus fumigatus	Fungus	Unassigned	Aero Fungi	Aspergillus Endo-chitosanase	238	Q875I9.1	74629604	16

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Aspergillus fumigatus Af293	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 2	304	EAL89830.1	66849502	7
Aspergillus fumigatus Af293	Fungus	Unassigned	Aero Fungi	Aspergillus Endo-chitosanase	242	Q4WB37.1	74666748	16
Aspergillus fumigatus var. RP-2014	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 4	322	KEY81716.1	666434194	16
Aspergillus fumigatus var. RP-2014	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 7	270	KEY78748.1	666431137	16
Aspergillus niger	Fungus	Unassigned	Aero Fungi	Aspergillus Asp f 18 and Asp n 18	533	AAA32702.1	289172	7
Aspergillus niger	Fungus	Asp n 14	Aero Fungi	Aspergillus Asp n 14	804	CAB06417.1	2181180	7
Aspergillus niger	Fungus	Asp n 14	Aero Fungi	Aspergillus Asp n 14	804	AAD13106.1	4235093	7
Aspergillus oryzae	Fungus	Asp o 21	Aero Fungi	Aspergillus Asp o 21	499	P0C1B3.1	94706935	7
Aspergillus oryzae	Fungus	Asp o 21.0101	Aero Fungi	Aspergillus Asp o 21	499	AAA32708.1	166531	15
Aspergillus oryzae	Fungus	Asp o 13	Aero Fungi	Aspergillus Oryzin Asp o 13, fl 13	403	P12547.2	129235	7
Aspergillus versicolor	Fungus	Unassigned	Aero Fungi	Aspergillus versicolor serine protease	403	ADE74975.1	294441150	16
Bacillus lentus	Bacteria	Unassigned	Bacteria airway	Bacillus lentus subtilisin	269	P29600.1	267048	9
Bacillus licheniformis	Bacteria	Unassigned	Bacteria airway	Bacillus licheniformis subtlilisin	379	P00780.1	135016	9
Bacillus licheniformis	Bacteria	Unassigned	Bacteria airway	Bacillus licheniformis subtlilisin	374	AAG31026.1	11127680	9
Bacillus sp.	Bacteria	Unassigned	Bacteria airway	Bacillus lentus Esperase	361	BAA05540.1	1225905	9
Balanus rostratus	Crustacean	Unassigned	Food Animal	Balanus r tropomyosin	284	BAF46896.1	125659386	9
Bassia scoparia	summer cypress	Unassigned	Aero Plant	Bassia scoparia	Koc s 1	167	AKV72169.1	914410012

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Bassia scoparia	summer cypress	Unassigned	Unassigned	Kochia scoparia Koc s 2.01	133	AIV43661.1	701225194	17
Batillus cornutus	Japanese turban shell	Unassigned	Food Animal	Batillus Tur c1	284	BAH10149.1	219806588	10
Bertholletia excelsa	Brazil nut	Ber e 1	Food Plant	Bertholletia Ber e 1	146	P04403.2	112754	7
Bertholletia excelsa	Brazil nut	Ber e 2	Food Plant	Bertholletia Ber e 2	465	AAO38859.1	30313867	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	51	A45786	320545	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA54696.1	534898	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	159	CAA54695.1	534900	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA54694.1	534910	7
Betula pendula	European white birch	Bet v 1.1601	Aero Plant	Betula Bet v 1	160	CAA96546.1	1321714	7
Betula pendula	European white birch	Bet v 1.1701	Aero Plant	Betula Bet v 1	160	CAA96539.1	1321716	7
Betula pendula	European white birch	Bet v 1.1801	Aero Plant	Betula Bet v 1	160	CAA96540.1	1321718	7
Betula pendula	European white birch	Bet v 1.1502	Aero Plant	Betula Bet v 1	160	CAA96541.1	1321720	7
Betula pendula	European white birch	Bet v 1.1901	Aero Plant	Betula Bet v 1	160	CAA96542.1	1321722	7
Betula pendula	European white birch	Bet v 1.2001	Aero Plant	Betula Bet v 1	160	CAA96543.1	1321724	7
Betula pendula	European white birch	Bet v 1.2101	Aero Plant	Betula Bet v 1	160	CAA96544.1	1321726	7
Betula pendula	European white birch	Bet v 1.2201	Aero Plant	Betula Bet v 1	160	CAA96547.1	1321728	7
Betula pendula	European white birch	Bet v 1m/n	Aero Plant	Betula Bet v 1	160	P43186.2	1168710	7
Betula pendula	European white birch	Bet v 1.0108	Aero Plant	Betula Bet v 1	160	CAB02155.1	1542861	7
Betula pendula	European white birch	Bet v 1.0109	Aero Plant	Betula Bet v 1	160	CAB02156.1	1542863	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Betula pendula	European white birch	Bet v 1.0110	Aero Plant	Betula Bet v 1	160	CAB02157.1	1542865	7
Betula pendula	European white birch	Bet v 1.0111	Aero Plant	Betula Bet v 1	160	CAB02158.1	1542867	7
Betula pendula	European white birch	Bet v 1.0112	Aero Plant	Betula Bet v 1	160	CAB02159.1	1542869	7
Betula pendula	European white birch	Bet v 1.0113	Aero Plant	Betula Bet v 1	160	CAB02160.1	1542871	7
Betula pendula	European white birch	Bet v 1.0114	Aero Plant	Betula Bet v 1	160	CAB02161.1	1542873	7
Betula pendula	European white birch	Bet v 1.2301	Aero Plant	Betula Bet v 1	160	CAA96545.1	2414158	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA05186.1	2564220	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA05187.1	2564222	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA05188.1	2564224	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA05190.1	2564228	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA07318.1	4006928	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA07319.1	4006945	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA07323.1	4006953	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA07324.1	4006955	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA07325.1	4006957	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA07326.1	4006959	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA07327.1	4006961	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA07329.1	4006965	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	160	CAA07330.1	4006967	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	159	CAA04823.1	4376216	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	159	CAA04826.1	4376219	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	159	CAA04827.1	4376220	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	159	CAA04828.1	4376221	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	159	CAA04829.1	4376222	7
Betula pendula	European white birch	Bet v 1 b1	Aero Plant	Betula Bet v 1	160	AAD26560.1	4590392	7
Betula pendula	European white birch	Bet v 1 b2	Aero Plant	Betula Bet v 1	160	AAD26561.1	4590394	7
Betula pendula	European white birch	bet v 1 b3	Aero Plant	Betula Bet v 1	160	AAD26562.1	4590396	7
Betula pendula	European white birch	Bet v 1.0701	Aero Plant	Betula Bet v 1	160	P43180.2	1168706	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	159	1QMR_A	11514622	7
Betula pendula	European white birch	Bet v 1x	Aero Plant	Betula Bet v 1	21	AAP37482.1	30908931	7
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	159	1LLT_A	38492423	7
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 1	43	AAB20452.1	239734	7
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 1	120	CAA07328.1	4006963	7
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 1	120	CAA07320.1	4006947	7
Betula pendula	European white birch	Bet v 1.0203	Aero Plant	Betula Bet v 1	160	CAA54488.1	452742	8
Betula pendula	European white birch	Bet v 1	Aero Plant	Betula Bet v 1	159	1B6F_A	159162097	9
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 1	159	4BK7_A	560188693	15
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 1	159	4B9R_A	550544347	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 1	160	4BKC_A	565807648	15
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 1	159	4BKD_A	560188694	15
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 1	159	4BK6_B	560188692	15
Betula pendula	European white birch	Bet v 1.0101	Aero Plant	Betula Bet v 1	160	CAA33887.1	17938	15
Betula pendula	European white birch	Bet v 1.0102	Aero Plant	Betula Bet v 1	160	CAA54482.1	452732	15
Betula pendula	European white birch	Bet v 1.0103	Aero Plant	Betula Bet v 1	160	CAA54483.1	452734	15
Betula pendula	European white birch	Bet v 1.0104	Aero Plant	Betula Bet v 1	160	CAA54484.1	452736	15
Betula pendula	European white birch	Bet v 1.0106	Aero Plant	Betula Bet v 1	160	CAA54487.1	452740	15
Betula pendula	European white birch	Bet v 1.0107	Aero Plant	Betula Bet v 1	160	CAA54489.1	452744	15
Betula pendula	European white birch	Bet v 1.0201	Aero Plant	Betula Bet v 1	160	CAA54421.1	450885	15
Betula pendula	European white birch	Bet v 1.0202	Aero Plant	Betula Bet v 1	160	CAA54481.1	452730	15
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 1	159	4BTZ_A	661918055	16
Betula pendula	European white birch	Unassigned	Unassigned	Betula Bet v 1	159	4Z3L_D	955264732	17
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 1b	51	B45786	320546	7
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 2	133	1CQA_A	157830684	9
Betula pendula	European white birch	Bet v 2.0101	Aero Plant	Betula Bet v 2	133	AAA16522.1	166953	11
Betula pendula	European white birch	Unassigned	Aero Plant	Betula Bet v 2	133	A4K9Z8.1	576017922	15
Betula pendula	European white birch	Bet v 3.0101	Aero Plant	Betula Bet v 3	205	CAA55854.1	488605	15
Betula pendula	European white birch	Bet v 4.0101	Aero Plant	Betula Bet v 4	85	CAA60628.1	809536	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Betula pendula	European white birch	Bet v 6.0102	Aero Plant	Betula Bet v 6	308	AAG22740.1	10764491	7
Betula pendula	European white birch	Bet v 7	Aero Plant	Betula Bet v 7	173	CAC84116.1	21886603	7
Betula pendula	European white birch	Bet v 8.0101	Aero Plant	Betula Bet v 8 glutathione S-transferase	237	AHF71027.1	573005958	16
Betula platyphylla	Japanese white birch	Unassigned	Aero Plant	Betula Bet v 1	160	BAB21489.1	12583681	7
Betula platyphylla	Japanese white birch	Unassigned	Aero Plant	Betula Bet v 1	160	BAB21490.1	12583683	7
Betula platyphylla	Japanese white birch	Unassigned	Aero Plant	Betula Bet v 1	160	BAB21491.1	12583685	7
Betula sp.	Birch	Unassigned	Aero Plant	Betula Bet v 1	51	AAB25850.1	298736	7
Betula sp.	Birch	Unassigned	Aero Plant	Betula Bet v 1b	51	AAB25851.1	298737	7
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella chymotrypsin-like	252	AJO53282.1	757943154	16
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella 36 kDa allergen	20	AAB29344.1	544618	7
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella 36 kDa allergen	25	AAB29345.1	544619	7
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella arginine kinase Bla g 9	356	ACM24358.1	221602737	10
Blattella germanica	German cockroach	Bla g 9.0101	Unassigned	Blattella arginine kinase Bla g 9	356	ABC86902.1	86160922	18
Blattella germanica	German cockroach	Bla g 1.0201	Aero Insect	Blattella Bla g 1	492	AAD13531.1	4240395	7
Blattella germanica	German cockroach	Bla g 1.0101	Aero Insect	Blattella Bla g 1	412	AAD13530.2	4572592	7
Blattella germanica	German cockroach	Bla g 11.0101	Aero Insect	Blattella Bla g 11 alpha Amylase	515	ABC68516.1	85002763	15
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella Bla g 2	330	1YG9_A	62738637	7
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella Bla g 2	352	ABP35603.1	145105726	9
Blattella germanica	German cockroach	Bla g 2.0101	Aero Insect	Blattella Bla g 2	352	AAA86744.1	1176397	11

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella Bla g 2	334	3LIZ_A	315113421	12
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella Bla g 3	657	ACY40650.1	262272875	11
Blattella germanica	German cockroach	Bla g 3.0101	Aero Insect	Blattella Bla g 3	657	ACY40651.1	262272877	11
Blattella germanica	German cockroach	Bla g 4	Aero Insect	Blattella Bla g 4	182	AAA87851.1	1166573	7
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella Bla g 4	182	ABP04043.1	144952778	9
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella Bla g 4	181	ACJ37389.1	212675308	10
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella Bla g 4	191	ACF53836.1	194350815	11
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella Bla g 4	190	ACF53837.1	194350817	11
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella Bla g 5	200	ABP04044.1	144952780	9
Blattella germanica	German cockroach	Bla g 5.0101	Aero Insect	Blattella Bla g 5	200	AAB72147.1	2326190	11
Blattella germanica	German cockroach	Bla g 6.0101	Aero Insect	Blattella Bla g 6	151	ABB89296.1	82704032	8
Blattella germanica	German cockroach	Bla g 6.0201	Aero Insect	Blattella Bla g 6	151	ABB89297.1	82704034	8
Blattella germanica	German cockroach	Bla g 6.0301	Aero Insect	Blattella Bla g 6	154	ABB89298.1	82704036	8
Blattella germanica	German cockroach	Bla g 7.0101	Aero Insect	Blattella Bla g 7	284	AAF72534.1	8101069	7
Blattella germanica	German cockroach	Unassigned	Aero Insect	Blattella delta GST	216	ABX57814.1	161137518	11
Blomia tropicalis	Mite	Blo t 1	Aero Mite	Blomia Blo t 1.01	221	AAK58415.1	14276828	7
Blomia tropicalis	Mite	Unassigned	Aero Mite	Blomia Blo t 1.02	333	AAQ24541.1	33667928	8
Blomia tropicalis	Mite	Unassigned	Aero Mite	Blomia Blo t 1.02	333		2	8
Blomia tropicalis	Mite	Blo t 10.0101	Aero Mite	Blomia Blo t 10	284	ABU97466.1	156938889	9

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Blomia tropicalis	Mite	Blo t 11	Aero Mite	Blomia Blo t 11	875	AAM83103.1	21954740	7
Blomia tropicalis	Mite	Blo t 12	Aero Mite	Blomia Blo t 12	144	AAA78904.1	902012	7
Blomia tropicalis	Mite	Unassigned	Aero Insect	Blomia Blo t 12	69	2MFK_A	723586656	16
Blomia tropicalis	Mite	Blo t 13	Aero Mite	Blomia Blo t 13.01	130	AAC80579.1	1377859	7
Blomia tropicalis	Mite	Unassigned	Aero Mite	Blomia Blo t 21	129	ABH06350.1	111120432	8
Blomia tropicalis	Mite	Unassigned	Aero Mite	Blomia Blo t 21	129	ABH06347.1	111494253	8
Blomia tropicalis	Mite	Unassigned	Aero Mite	Blomia Blo t 21	129	ABH06346.1	111120424	8
Blomia tropicalis	Mite	Unassigned	Aero Mite	Blomia Blo t 21	129	ABH06348.1	111120428	8
Blomia tropicalis	Mite	Blo t 21.0101	Aero Insect	Blomia Blo t 21	129	AAX34047.1	60679570	9
Blomia tropicalis	Mite	Blo t 3.0101	Aero Mite	Blomia Blo t 3	266	AAM10779.1	25989482	7
Blomia tropicalis	Mite	Unassigned	Aero Mite	Blomia Blo t 3	266	AAQ24542.1	33667930	8
Blomia tropicalis	Mite	Unassigned	Aero Insect	Blomia Blo t 4 alpha amylase	506	AAQ24543.1	33667932	8
Blomia tropicalis	Mite	Blo t 5	Aero Mite	Blomia Blo t 5	134	AAD10850.1	4204917	7
Blomia tropicalis	Mite	Unassigned	Aero Mite	Blomia Blo t 5	134	ABH06352.1	111120436	9
Blomia tropicalis	Mite	Unassigned	Aero Mite	Blomia Blo t 5	134	ABH06359.1	111120450	9
Blomia tropicalis	Mite	Unassigned	Aero Mite	Blomia Blo t 5	119	2JMH_A	160285626	9
Blomia tropicalis	Mite	Unassigned	Unassigned	Blomia Blo t 5	135	APU87558.1	1131385191	18
Blomia tropicalis	Mite	Unassigned	Unassigned	Blomia Blo t 5	133	APU87557.1	1131385189	18
Blomia tropicalis	Mite	Unassigned	Unassigned	Blomia Blo t 5	135	APU87556.1	1131385187	18

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Blomia tropicalis	Mite	Unassigned	Unassigned	Blomia Blo t 5	134	APU87554.1	1131385183	18
Blomia tropicalis	Mite	Unassigned	Aero Insect	Blomia Blo t 7	192	AAQ24545.1	33667936	8
Blomia tropicalis	Mite	Blo t 7.0101	Unassigned	Blomia Blo t 7	195	ASX95438.1	1241067909	18
Blomia tropicalis	Mite	Unassigned	Aero Insect	Blomia Blo t 8	236	AAP35069.1	37958149	8
Blomia tropicalis	Mite	Blo t 8.0101	Aero Insect	Blomia Blo t 8	236	ACV04860.1	256665455	11
Bombus pennsylvanicus	Bumblebee	Bom p 1.0101	Venom or Salivary	Bombus Bom p 1	136	Q7M4I6.1	47117013	12
Bombus pennsylvanicus	Bumblebee	Bom p 4.0101	Venom or Salivary	Bombus Bom p 4 protease	243	Q7M4I3.1	75009997	12
Bombus terrestris	Bumblebee	Bom t 1.0101	Venom or Salivary	Bombus Bom t 1	136	P82971.1	14423832	7
Bombus terrestris	Bumblebee	Unassigned	Venom or Salivary	Bombus Bom t 4 protease	20	P0CH88.1	313471465	12
Bombyx mori	Silkworm	Bomb m 1.0101	Aero Insect	Bombyx Bomb m 1	355	ABB88514.1	82658675	15
Bos grunniens mutus	Yak	Unassigned	Food Animal	Bos Bos d 11 beta casein	259	XP_005902099.2	942073448	16
Bos taurus	Bovine	Unassigned	Food Animal	Bos Alpha-s1 casein	93	AAA62707.1	162650	7
Bos taurus	Bovine	Unassigned	Food Animal	Bos Alpha-s1 casein	214	AAA30429.1	162794	7
Bos taurus	Bovine	Unassigned	Food Animal	Bos Alpha-s1 casein	76	AAA30478.1	162927	7
Bos taurus	Bovine	Bos d 9.0101	Food Animal	Bos Alpha-s1 casein	214	NP_851372.1	30794348	8
Bos taurus	Bovine	Unassigned	Food Animal	Bos Alpha-s1 casein	205	ABW98943.1	159793197	9
Bos taurus	Bovine	Unassigned	Food Animal	Bos Alpha-s1 casein	172	ABW98945.1	159793201	9
Bos taurus	Bovine	Unassigned	Food Animal	Bos Alpha-s1 casein	129	ABW98953.1	159793217	9
Bos taurus	Bovine	Bos d 10.0101	Food Animal	Bos Bos d 10	222	NP_776953.1	27806963	15
Bos taurus	Bovine	Unassigned	Food Animal	Bos Bos d 11 beta casein	224	AAA30430.1	162797	7
Bos taurus	Bovine	Unassigned	Food Animal	Bos Bos d 11 beta casein	224	AAA30431.1	162805	7
Bos taurus	Bovine	Unassigned	Food Animal	Bos Bos d 11 beta casein	224	AAB29137.1	459292	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Bos taurus	Bovine	Unassigned	Food Animal	Bos Bos d 12	190	AAA30433.1	162811	7
Bos taurus	Bovine	Bos d 12.0101	Food Animal	Bos Bos d 12	190	NP_776719.1	27881412	15
Bos taurus	Bovine	Unassigned	Aero Animal	Bos Bos d 2	172	Q28133.1	2497701	9
Bos taurus	Bovine	Bos d 3	Aero Animal	Bos Bos d 3	101	Q28050.1	2493414	7
Bos taurus	Bovine	Unassigned	Food Animal	Bos Bos d 4	142	CAA29664.1	295774	7
Bos taurus	Bovine	Bos d 4.0101	Food Animal	Bos Bos d 4	142	AAA30615.1	163283	15
Bos taurus	Bovine	Bos d 5	Food Animal	Bos Bos d 5	178	CAA32835.1	520	7
Bos taurus	Bovine	Unassigned	Food Animal	Bos Bos d 5	14	AAA30413.1	162750	7
Bos taurus	Bovine	Unassigned	Food Animal	Bos Bos d 5	178	P02754.3	125910	9
Bos taurus	Bovine	Unassigned	Food Animal	Bos Bos d 5	178	ACG59280.1	195957138	10
Bos taurus	Bovine	Bos d 6	Food Animal	Bos Bos d 6	607	AAA51411.1	162648	7
Bos taurus	Bovine	Unassigned	Food Animal	Bos Bos d 6	607	CAA76847.1	3336842	7
Bos taurus	Bovine	Unassigned	Vaccine	Bos collagen alpha2	1364	NP_776945.1	27806257	11
Bos taurus	Bovine	Unassigned	Food Animal	Bos lactotransferrin	708	NP_851341.1	30794292	8
Brassica juncea	Mustard	Bra j 1	Food Plant	Brassica Bra j 1 2S albumin	129	P80207.1	32363444	9
Brassica napus	Rape	Bra n 1	Food Plant	Bra n 1	125	P80208.1	75107016	9
Brassica napus	Rape	Unassigned	Aero Plant	Bra n Bra r 2	83	S65144	2129801	7
Brassica napus	Rape	Unassigned	Aero Plant	Bra n Bra r 2	83	S65145	2129802	7
Brassica napus	Rape	Unassigned	Food Plant	Brassica napus 2S albumin	109	AAN86249.1	26985163	7
Brassica oleracea var. oleracea	Wild cabbage	Unassigned	Aero Plant	Brassica Bra o 3 LTP full length	112	XP_013623213.1	922434456	16
Brassica rapa	Turnip	Unassigned	Aero Plant	Bra n Bra r 2	80	S65143	2129805	7
Brassica rapa	Turnip	Bra r 1.0101	Food Plant	Brassica Bra r 1	178	CAA46782.1	17697	9
Brassica rapa	Turnip	Bra r 5.0101	Food Plant	Brassica Calcim binding protein Group I	79	BAA09634.1	1255540	15
Brassica rapa subsp. rapa	Turnip	Unassigned	Aero Plant	Bra n Bra r 2	83	P69199.1	59800146	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Brassica rapa subsp. rapa	Turnip	Bra r 2.0101	Contact	Brassica Bra r 2	91	P81729.1	32363456	9
Candida albicans	Yeast	Cand a 1.0101	Contact	Candida Cand a 1 Alcohol dehydrogenase	350	CAA57342.1	608690	15
Candida albicans	Yeast	Cand a 3.0101	Contact	Candida Cand a 3 Peroxysomal protein	236	AAN11300.1	37548637	7
Candida albicans	Yeast	Unassigned	Contact	Candida Enolase 1	440	P30575.1	232054	7
Canis familiaris	Dog	Can f 1.0101	Aero Animal	Canis Can f 1 Lipocalin	174	AAC48794.1	2598974	11
Canis familiaris	Dog	Can f 2	Aero Animal	Canis Can f 2 Lipocalin	177	CAD82911.1	29292272	7
Canis familiaris	Dog	Can f 2	Aero Animal	Canis Can f 2 Lipocalin	179	CAD82912.1	29292274	7
Canis familiaris	Dog	Can f 2.0101	Aero Animal	Canis Can f 2 Lipocalin	180	AAC48795.1	2598976	11
Canis familiaris	Dog	Can f 3	Aero Animal	Canis Can f 3 Serum albumin	265	AAB30434.1	633938	7
Canis familiaris	Dog	Can f 3	Aero Animal	Canis Can f 3 Serum albumin	585	CAA76841.1	3319897	7
Canis familiaris	Dog	Can f 3.0101	Aero Animal	Canis Can f 3 Serum albumin	608	BAC10663.1	22531688	15
Canis familiaris	Dog	Can f 4.0101	Aero Animal	Canis Can f 4 epithelial 18 kDa	174	ACY38525.1	262232390	12
Canis familiaris	Dog	Unassigned	Aero Animal	Canis Can f 4 epithelial 18 kDa	174	AHY24648.1	625295108	16
Canis familiaris	Dog	Can f 5.0101	Aero Animal	Canis Can f 5	260	CAA68720.1	868	15
Canis familiaris	Dog	Can f 6.0101	Aero Animal	Canis Can f 6 Lipocalin	190	CCF72371.1	374092884	13
Canis familiaris	Dog	Unassigned	Unassigned	Canis familiaris	Can f 7	149	AAB34263.1	945179
Cannabis sativa	Hemp	Can s 3.0101	Aero Plant	Cannabis LTP Can s 3	91	CCK33472.1	571256597	15
Capsicum annuum	Bell pepper	Cap a 1	Food Plant	Capsicum Cap a 1	246	CAC34055.2	16609959	7
Capsicum annuum	Bell pepper	Cap a 2	Food Plant	Capsicum Cap a 2	131	CAD10376.1	16555785	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Carica papaya	Papaya	Car p 1	Aero Plant	Carica Car p 1 not official name	345	AAB02650.1	167391	7
Carpinus betulus	Hornbeam	Car b 1.0102	Aero Plant	Carpinus Car b 1	159	CAA47357.1	402745	7
Carpinus betulus	Hornbeam	Car b 1.0103	Aero Plant	Carpinus Car b 1	160	CAB02206.1	1545875	7
Carpinus betulus	Hornbeam	Car b 1.0104	Aero Plant	Carpinus Car b 1	160	CAB02207.1	1545877	7
Carpinus betulus	Hornbeam	Car b 1.0105	Aero Plant	Carpinus Car b 1	160	CAB02208.1	1545879	7
Carpinus betulus	Hornbeam	Car b 1.0108	Aero Plant	Carpinus Car b 1	160	CAB02215.1	1545893	7
Carpinus betulus	Hornbeam	Car b 1.0301	Aero Plant	Carpinus Car b 1	161	CAB02216.1	1545895	7
Carpinus betulus	Hornbeam	Car b 1.0302	Aero Plant	Carpinus Car b 1	161	CAB02217.1	1545897	7
Carpinus betulus	Hornbeam	Unassigned	Aero Plant	Carpinus Car b 1	40	AAB20453.1	239735	7
Carpinus betulus	Hornbeam	Car b 1.0113	Aero Plant	Carpinus Car b 1	160	ABZ81044.1	167472845	10
Carpinus betulus	Hornbeam	Car b 1.0109	Aero Plant	Carpinus Car b 1	160	ABZ81040.1	167472837	10
Carpinus betulus	Hornbeam	Car b 1.0112	Aero Plant	Carpinus Car b 1	160	ABZ81043.1	167472843	10
Carpinus betulus	Hornbeam	Car b 1.0111	Aero Plant	Carpinus Car b 1	160	ABZ81042.1	167472841	10
Carpinus betulus	Hornbeam	Car b 1.0110	Aero Plant	Carpinus Car b 1	160	ABZ81041.1	167472839	10
Carpinus betulus	Hornbeam	Unassigned	Aero Plant	Carpinus Car b 1	80	AAB34907.1	1008578	12
Carpinus betulus	Hornbeam	Unassigned	Aero Plant	Carpinus Car b 1	80	AAB34908.1	1008579	12
Carpinus betulus	Hornbeam	Unassigned	Aero Plant	Carpinus Car b 1	80	AAB34909.1	1008580	12
Carpinus betulus	Hornbeam	Car b 1.0101	Aero Plant	Carpinus Car b 1	159	CAA47366.1	402743	15
Carpinus betulus	Hornbeam	Car b 1.0106	Aero Plant	Carpinus Car b 1	160	CAB02209.1	1545881	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Carpinus betulus	Hornbeam	Car b 1.0107	Aero Plant	Carpinus Car b 1	160	CAB02213.1	1545889	15
Carpinus betulus	Hornbeam	Car b 1.0201	Aero Plant	Carpinus Car b 1	159	CAA47367.1	402747	15
Carya illinoinensis	Pecan	Car i 1.0101	Food Plant	Carya Car i 1 Seed storage protein	143	AAO32314.1	28207731	7
Carya illinoinensis	Pecan	Car i 4.0101	Food Plant	Carya Car i 4 11s legumin	505	ABW86978.1	158998780	14
Carya illinoinensis	Pecan	Unassigned	Food Plant	Carya Car i 4 11s legumin	505	ABW86979.1	158998782	14
Carya illinoinensis	Pecan	Car i 2.0101	Food Plant	Carya illinoisnensis Car i 2 vicilin	792	ABV49590.1	157384600	15
Carya illinoinensis	Pecan	Unassigned	Food plant	Carya illinoisnensis Car i 2 vicilin	426	5E1R_F	1052244924	18
Caryota mitis	Fishtail Palm	Unassigned	Aero Plant	Caryota profilin	131	ABM53030.1	121277849	8
Castanea sativa	European chestnut	Cas s 1	Aero Plant	Castanea Cas s 1	160	CAD10374.1	16555781	7
Castanea sativa	European chestnut	Unassigned	Aero Plant	Castanea Cas s 1	159	ACJ23862.1	212291466	10
Castanea sativa	European chestnut	Cas s 1.0101	Aero Plant	Castanea Cas s 1	159	ACJ23861.1	212291464	10
Castanea sativa	European chestnut	Unassigned	Aero Plant	Castanea Cas s 1	159	ACJ23863.1	212291468	10
Castanea sativa	European chestnut	Cas s 5	Food Plant	Castanea Cas s 5	316	CAA64868.1	1359600	7
Castanea sativa	European chestnut	Unassigned	Food Plant	Castanea Cas s 5	298	ADN39439.1	307159110	12
Catharanthus roseus	Madagascar periwinkle	Unassigned	Aero Plant	Catharanthus cyclophilin	178	2MC9_A	659835152	16
Cavia porcellus	Domestic guinea pig	Cav p 1	Aero Animal	Cavia Cav p 1	15	P83507.1	32469617	7
Cavia porcellus	Domestic guinea pig	Cav p 2.0101	Aero Animal	Cavia Cav p 2	170	CAX62129.1	325910590	12
Cavia porcellus	Domestic guinea pig	Cav p 3.0101	Aero Animal	Cavia Cav p 3 lipocalin	170	CAX62130.1	325910592	12
Chamaecyparis obtusa	Japanese cypress	Cha o 1.0101	Aero Plant	Chamaecyparis Cha o 1	375	BAA08246.1	1514943	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Chamaecyparis obtusa	Japanese cypress	Unassigned	Aero Plant	Chamaecyparis Cha o 2	514	Q7M1E7.1	47606004	7
Chamaecyparis obtusa	Japanese cypress	Unassigned	Aero Plant	Chamaecyparis Cha o 2	419	BAF32143.1	114841683	8
Chamaecyparis obtusa	Japanese cypress	Unassigned	Aero Plant	Chamaecyparis obtusa Cha o 3	436	Manual_Cha_o_3	6	17
Charybdis feriatus	Crab	Cha f 1.0101	Food Animal	Charybdis Cha f 1	264	AAF35431.1	7024506	7
Chenopodium album	Pigweed	Che a 1	Aero Plant	Chenopodium Che a 1	168	AAL07319.1	22074346	7
Chenopodium album	Pigweed	Che a 2	Aero Plant	Chenopodium Che a 2	131	AAL92870.1	29465666	7
Chenopodium album	Pigweed	Unassigned	Aero Plant	Chenopodium Che a 2	133	ACR77509.1	238886048	11
Chenopodium album	Pigweed	Che a 3	Aero Plant	Chenopodium Che a 3	86	AAL92871.1	29465668	7
Chionoecetes opilio	Snow Crab	Unassigned	Food Animal	Chionoecetes tropomyosin	284	A2V735.1	308191588	12
Chironomus kiiensis	Midge	Chi k 10	Aero Insect	Chironomus Chi k 10	285	CAA09938.2	7321108	7
Chironomus thummi thummi	Midge	Chi t 1.01	Aero Insect	Chironomus Chi t 1	151	P02229.2	121219	7
Chironomus thummi thummi	Midge	Chi t 1.02	Aero Insect	Chironomus Chi t 1	151	P02230.1	121227	7
Chironomus thummi thummi	Midge	Chi t 2.0101	Aero Insect	Chironomus Chi t 2	158	P02221.2	2506460	7
Chironomus thummi thummi	Midge	Chi t 3.0601	Aero Insect	Chironomus Chi t 3	161	P84296.1	56405052	7
Chironomus thummi thummi	Midge	Chi t 3.0901	Aero Insect	Chironomus Chi t 3	151	P02227.1	121237	7
Chironomus thummi thummi	Midge	Chi t 3.0501	Aero Insect	Chironomus Chi t 3	161	P12548.1	121244	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Chironomus thummi thummi	Midge	Chi t 3.0701	Aero Insect	Chironomus Chi t 3	161	P84298.1	56405054	7
Chironomus thummi thummi	Midge	Chi t 3.0702	Aero Insect	Chironomus Chi t 3	161	P12549.1	121248	7
Chironomus thummi thummi	Midge	Chi t 3.0801	Aero Insect	Chironomus Chi t 3	162	P12550.1	121249	7
Chironomus thummi thummi	Midge	Chi t 3.0301	Aero Insect	Chironomus Chi t 3	161	P02226.2	56405306	7
Chironomus thummi thummi	Midge	Chi t 3.0101	Aero Insect	Chironomus Chi t 3	160	P02222.2	1707908	7
Chironomus thummi thummi	Midge	Chi t 3.0401	Aero Insect	Chironomus Chi t 3	161	P02223.2	1707911	7
Chironomus thummi thummi	Midge	Chi t 3.0201	Aero Insect	Chironomus Chi t 3	162	P02224.2	2506461	7
Chironomus thummi thummi	Midge	Chi t 4	Aero Insect	Chironomus Chi t 4	151	P02231.1	121256	7
Chironomus thummi thummi	Midge	Chi t 9	Aero Insect	Chironomus Chi t 9	151	P02228.1	121259	7
Citrullus lanatus	Watermellon	Citr l 2	Plant food	Citrullus lanatus Citr l 2	131	AAU43733.1	52352489	17
Citrus limon	Lemon	Cit l 3.0101	Food Plant	Citrus LTP Cit s 3/Cit l 3	20	P84160.1	52783176	7
Citrus sinensis	Navel orange	Cit s 1.0101	Food Plant	Citrus Cit s 1	25	P84159.1	52782810	7
Citrus sinensis	Navel orange	Cit s 2.0101	Food Plant	Citrus Cit s 2	131	CAI23765.1	56000996	7
Citrus sinensis	Navel orange	Unassigned	Food Plant	Citrus LTP Cit s 3/Cit l 3	20	P84161.1	52783177	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Citrus sinensis	Navel orange	Cit s 3	Food Plant	Citrus LTP Cit s 3/Cit l 3	91	CAH03799.1	50199132	7
Cladosporium cladosporioides	Fungus	Cla c 14.0101	Aero Fungi	Cladosporium Cla c 14	325	ADK47394.1	301015198	15
Cladosporium cladosporioides	Fungus	Cla c 9.0101	Aero Fungi	Cladosporium Cla c 9 Davidiella	388	ABQ59329.1	148361511	11
Clupea harengus	Atlantic herring	Clu h 1.0101	Food Animal	Clupea Clu h 1	109	CAQ72970.1	242253963	11
Clupea harengus	Atlantic herring	Clu h 1.0201	Food Animal	Clupea Clu h 1	110	CAQ72971.1	242253965	11
Clupea harengus	Atlantic herring	Clu h 1.0301	Food Animal	Clupea Clu h 1	109	CAQ72972.1	242253967	11
Cochliobolus lunatus	Fungus	Cur l 2.01	Aero Fungi	Curvularia lunata enolase Cur l 2.01 Cochliobolus	440	AAK67491.1	14585753	8
Cochliobolus lunatus	Fungus	Cur l 3.0101	Aero Fungi	Curvularia lunatua Cur l 3 Cochliobolus	108	AAK67492.1	14585755	15
Cochliobolus lunatus	Fungus	Cur l 4.0101	Aero Fungi	Curvularia Cur l 4	506	ACF19589.1	193507493	15
Cochliobolus lunatus	Fungus	Unassigned	Unassigned	Curvularia lunata alcohol dehydrogenase	352	ABC88428.1	86278351	17
Coffea arabica	Coffee	Cof a 3.0101	Food Plant	Coffea Cof a 3	65	AGL34968.1	494319676	15
Coffea arabica	Coffee	Cof a 1.0101	Food Plant	Coffea Cof a 1	263	ADH10372.1	296399179	15
Coffea arabica	Coffee	Cof a 2.0101	Food Plant	Coffea Cof a 2	80	AGL34967.1	494319674	15
Coprinus comatus	Shaggy mane	Cop c 1	Food Fungi	Coprinus Cop c 1	81	CAB39376.1	4538529	7
Corylus avellana	European hazelnut	Cor a 1.0103	Aero Plant	Corylus Cor a 1	160	CAA50325.1	22684	7
Corylus avellana	European hazelnut	Cor a 1.0104	Aero Plant	Corylus Cor a 1	160	CAA50326.1	22686	7
Corylus avellana	European hazelnut	Cor a 1.0102	Aero Plant	Corylus Cor a 1	160	CAA50328.1	22690	7
Corylus avellana	European hazelnut	Cor a 1.0201	Aero Plant	Corylus Cor a 1	160	CAA96548.1	1321731	7
Corylus avellana	European hazelnut	Cor a 1.0301	Aero Plant	Corylus Cor a 1	160	CAA96549.1	1321733	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Corylus avellana	European hazelnut	Cor a 1.0401	Food Plant	Corylus Cor a 1	161	AAD48405.1	5726304	7
Corylus avellana	European hazelnut	Cor a 1.0402	Food Plant	Corylus Cor a 1	161	AAG40329.1	11762102	7
Corylus avellana	European hazelnut	Cor a 1.0403	Food Plant	Corylus Cor a 1	161	AAG40330.1	11762104	7
Corylus avellana	European hazelnut	Cor a 1.0404	Food Plant	Corylus Cor a 1	161	AAG40331.1	11762106	7
Corylus avellana	European hazelnut	Cor a 1.0101	Food Plant	Corylus Cor a 1	160	CAA50327.1	22688	15
Corylus avellana	European hazelnut	Cor a 11	Food Plant	Corylus Cor a 11	448	AAL86739.1	19338630	7
Corylus avellana	European hazelnut	Cor a 12.0101	Food Plant	Corylus Cor a 12	159	AAO67349.2	49617323	15
Corylus avellana	European hazelnut	Cor a 13.0101	Food Plant	Corylus Cor a 13 Oleosin	140	AAO65960.1	29170509	7
Corylus avellana	European hazelnut	Cor a 14.0101	Food Plant	Corylus Cor a 14 2S albumin	147	ACO56333.1	226437844	11
Corylus avellana	European hazelnut	Cor a 2.0101	Aero Plant	Corylus Cor a 2 profilins	131	AAK01235.1	12659206	7
Corylus avellana	European hazelnut	Cor a 2.0102	Aero Plant	Corylus Cor a 2 profilins	131	AAK01236.1	12659208	7
Corylus avellana	European hazelnut	Unassigned	Food Plant	Corylus Cor a 2 profilins	131	A4KA41.1	576017879	15
Corylus avellana	European hazelnut	Unassigned	Food Plant	Corylus Cor a 2 profilins	133	A4KA40.1	576017878	15
Corylus avellana	European hazelnut	Unassigned	Food Plant	Corylus Cor a 2 profilins	133	A4KA44.1	576017819	15
Corylus avellana	European hazelnut	Unassigned	Food Plant	Corylus Cor a 2 profilins	131	A4KA43.1	576017779	15
Corylus avellana	European hazelnut	Unassigned	Food Plant	Corylus Cor a 2 profilins	133	A4KA45.1	576017777	15
Corylus avellana	European hazelnut	Unassigned	Food Plant	Corylus Cor a 2 profilins	133	A4KA39.1	576017776	15
Corylus avellana	European hazelnut	Cor a 8	Food Plant	Corylus Cor a 8	115	AAK28533.1	13507262	7
Corylus avellana	European hazelnut	Cor a 9	Food Plant	Corylus Cor a 9	515	AAL73404.1	18479082	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Corylus avellana	European hazelnut	Unassigned	Aero Plant	Corylus Cor a 9	514	AHA36627.1	557792009	16
Crangon crangon	Shrimp	Cra c 1.0101	Food Animal	Crangon Cra c 1 tropomyosin	284	ACR43473.1	238477263	12
Crangon crangon	Shrimp	Cra c 2.0101	Food Animal	Crangon Cra c 2 arginine kinase	356	ACR43474.1	238477265	12
Crangon crangon	Shrimp	Cra c 4.0101	Food Animal	Crangon Cra c 4 sarcoplasmic calcium-binding protein	193	ACR43475.1	238477327	12
Crangon crangon	Shrimp	Cra c 5.0101	Food Animal	Crangon Cra c 5 myosin light chain	153	ACR43477.1	238477331	12
Crangon crangon	Shrimp	Cra c 6.0101	Food Animal	Crangon Cra c 6 troponin C	150	ACR43478.1	238477333	12
Crangon crangon	Shrimp	Cra c 8.0101	Food Animal	Crangon Cra c 8 triosephosphate isomerase	249	ACR43476.1	238477329	12
Crassostrea gigas	American oyster	Cra g 1.0102	Food Animal	Crassostrea Tropomyosin Cra g 1	284	BAH10152.1	219806594	10
Crassostrea gigas	American oyster	Cra g 1.0101	Unassigned	Crassostrea Tropomyosin Cra g 1	284	ARX70262.1	1203820203	18
Crassostrea virginica	Eastern oyster	Unassigned	Food Animal	Crassostrea Tropomyosin Cra g 1	160	AAC61869.1	3668408	7
Crocus sativus	Saffron crocus	Cro s 2.0101	Aero Plant	Crocus profilin Cro s 2	131	AAW81034.1	58700651	7
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria class IV chitinase	281	BAD77932.1	56550550	7
Cryptomeria japonica	Japanese cedar	Cry j 1.0102	Aero Plant	Cryptomeria Cry j 1	374	BAA05543.1	493634	8
Cryptomeria japonica	Japanese cedar	Cry j 1.0101	Aero Plant	Cryptomeria Cry j 1	374	BAA05542.1	493632	15
Cryptomeria japonica	Japanese cedar	Cry j 1.0103	Aero Plant	Cryptomeria Cry j 1	374	BAA07020.1	516728	15
Cryptomeria japonica	Japanese cedar	Cry j 2	Aero Plant	Cryptomeria Cry j 2	514	P43212.1	1171004	7
Cryptomeria japonica	Japanese cedar	Cry j 2	Aero Plant	Cryptomeria Cry j 2	514	BAC23082.1	24898904	7
Cryptomeria japonica	Japanese cedar	Cry j 2	Aero Plant	Cryptomeria Cry j 2	514	BAC23083.1	24898906	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Cryptomeria japonica	Japanese cedar	Cry j 2	Aero Plant	Cryptomeria Cry j 2	514	BAC23084.1	24898908	7
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Cry j 2	514	BAF32105.1	114841607	8
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Cry j 2	514	BAF32110.1	114841617	8
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Cry j 2	514	BAF32116.1	114841629	8
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Cry j 2	514	BAF32119.1	114841635	8
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Cry j 2	514	BAF32122.1	114841641	8
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Cry j 2	514	BAF32128.1	114841653	8
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Cry j 2	514	BAF32130.1	114841657	8
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Cry j 2	514	BAF32133.1	114841663	8
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Cry j 2	514	BAF32134.1	114841665	8
Cryptomeria japonica	Japanese cedar	Cry j 2.0101	Aero Plant	Cryptomeria Cry j 2	514	BAA06172.1	506858	9
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Cry j 2	65	BAF45320.1	123299282	9
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria Isoflavone reductase-like protein	306	AAK27264.1	19847822	7
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria pollen allergen CJP-8	165	BAI94503.1	291621332	12
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria pollen allergen CPA63	472	BAJ04354.1	293329689	12
Cryptomeria japonica	Japanese cedar	Unassigned	Aero Plant	Cryptomeria thaumatin like Cry j 3.8	225	BAF51970.1	139002766	8
Cucumis melo	Muskmelon	Cuc m 1	Food Plant	Cucumis Cuc m 1	731	BAA06905.1	807698	7
Cucumis melo	Muskmelon	Unassigned	Food Plant	Cucumis Cuc m 2	131	CAD92666.1	31559374	7
Cucumis melo	Muskmelon	Cuc m 2	Food Plant	Cucumis Cuc m 2	131	AAW69549.1	58263793	7
Cucumis melo	Muskmelon	Cuc m 3.0101	Food Plant	Cucumis Cuc m 3	41	P83834.1	46396595	9

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Cucumis melo var. inodorus	Muskmelon	Unassigned	Food Plant	Cucumis Cuc m 3	151	ACB45874.1	171464770	9
Cucumis melo var. reticulatus	Netted muskmelon	Unassigned	Food Plant	Cucumis Cuc m 2	131	AAP13533.2	57021110	7
Cupressus arizonica	Arizona Cypress	Cup a 1	Aero Plant	Cupressus Cup a 1/Cup s 1	346	CAB62551.1	6562326	7
Cupressus arizonica	Arizona Cypress	Cup a 1	Aero Plant	Cupressus Cup a 1/Cup s 1	367	CAC37790.2	19069497	7
Cupressus arizonica	Arizona Cypress	Unassigned	Aero Plant	Cupressus Cup a 1/Cup s 1	347	ABK78766.1	118197955	8
Cupressus arizonica	Arizona Cypress	Unassigned	Aero Plant	Cupressus Cup a 4	165	ACY01951.1	261865475	11
Cupressus arizonica	Arizona Cypress	Unassigned	Aero Plant	Cupressus Cup s 3	199	CAC05258.1	9929163	7
Cupressus sempervirens	Mediterranean Cypress	Cup s 1.0101	Aero Plant	Cupressus Cup a 1/Cup s 1	367	AAF72625.1	8101711	7
Cupressus sempervirens	Mediterranean Cypress	Cup s 1.0102	Aero Plant	Cupressus Cup a 1/Cup s 1	367	AAF72626.1	8101713	7
Cupressus sempervirens	Mediterranean Cypress	Cup s 1.0103	Aero Plant	Cupressus Cup a 1/Cup s 1	367	AAF72627.1	8101715	7
Cupressus sempervirens	Mediterranean Cypress	Cup s 1.0104	Aero Plant	Cupressus Cup a 1/Cup s 1	367	AAF72628.1	8101717	7
Cupressus sempervirens	Mediterranean Cypress	Cup s 1.0105	Aero Plant	Cupressus Cup a 1/Cup s 1	367	AAF72629.1	8101719	7
Cupressus sempervirens	Mediterranean Cypress	Cup s 3.0102	Aero Plant	Cupressus Cup s 3	225	AAR21074.1	38456228	7
Cupressus sempervirens	Mediterranean Cypress	Cup s 3.0101	Aero Plant	Cupressus Cup s 3	225	AAR21073.1	38456226	11
Cynodon dactylon	Bermuda grass	Cyn d 1	Aero Plant	Cynodon Cyn d 1	25	AAB28566.1	451274	7
Cynodon dactylon	Bermuda grass	Cyn d 1	Aero Plant	Cynodon Cyn d 1	38	AAB28567.1	451275	7
Cynodon dactylon	Bermuda grass	Cyn d 1	Aero Plant	Cynodon Cyn d 1	34	AAB32317.1	691726	7
Cynodon dactylon	Bermuda grass	Cyn d 1.0204	Aero Plant	Cynodon Cyn d 1	244	AAF80379.2	10314021	7
Cynodon dactylon	Bermuda grass	Cyn d 1.0201	Aero Plant	Cynodon Cyn d 1	244	AAK96255.1	15384338	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Cynodon dactylon	Bermuda grass	Cyn d 1.0202	Aero Plant	Cynodon Cyn d 1	262	AAL14077.1	16076693	7
Cynodon dactylon	Bermuda grass	Cyn d 1	Aero Plant	Cynodon Cyn d 1	262	AAL14078.1	16076695	7
Cynodon dactylon	Bermuda grass	Cyn d 1.0203	Aero Plant	Cynodon Cyn d 1	262	AAL14079.1	16076697	7
Cynodon dactylon	Bermuda grass	Cyn d 1.0101	Aero Plant	Cynodon Cyn d 1	246	AAB50734.2	7687901	10
Cynodon dactylon	Bermuda grass	Cyn d 12	Aero Plant	Cynodon Cyn d 12	131	CAA69670.1	2154730	7
Cynodon dactylon	Bermuda grass	Unassigned	Aero Plant	Cynodon Cyn d 7	71	CAA01909.1	1247373	7
Cynodon dactylon	Bermuda grass	Unassigned	Aero Plant	Cynodon Cyn d 7	73	CAA01910.1	1247375	7
Cynodon dactylon	Bermuda grass	Cyn d 7	Aero Plant	Cynodon Cyn d 7	82	CAA62634.1	1871507	7
Cynodon dactylon	Bermuda grass	Unassigned	Aero Plant	Cynodon Group 4 like-allergen FAD-linked oxidoredu	522	AAS02108.1	41393750	7
Cyprinus carpio	Carp	Cyp c 1.0101	Food Animal	Cyprinus Cyp c 1 Parvalbumin	109	CAC83658.1	17977825	7
Cyprinus carpio	Carp	Cyp c 1.0201	Food Animal	Cyprinus Cyp c 1 Parvalbumin	109	CAC83659.1	17977827	7
Dactylis glomerata	Orchard grass	Dac g 1	Aero Plant	Dactylis Dac g 1	264	CAD20406.1	18093991	7
Dactylis glomerata	Orchard grass	Dac g 1.0101	Aero Plant	Dactylis Dac g 1	240	AAP96759.1	33149333	7
Dactylis glomerata	Orchard grass	Dac g 2	Aero Plant	Dactylis Dac g 2	196	2103117A	1093120	7
Dactylis glomerata	Orchard grass	Dac g 2	Aero Plant	Dactylis Dac g 2	122	CAA10345.1	4007040	7
Dactylis glomerata	Orchard grass	Dac g 3	Aero Plant	Dactylis Dac g 3	96	AAB42200.1	1825459	7
Dactylis glomerata	Orchard grass	Dac g 4.0101	Aero Plant	Dactylis Dac g 4	55	P82946.1	32363463	9
Dactylis glomerata	Orchard grass	Dac g 5	Aero Plant	Dactylis Dac g 5	290	AAK62278.1	14423124	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dactylis glomerata	Orchard grass	Dac g 5	Aero Plant	Dactylis Dac g 5	265	CAD20405.1	18093971	7
Daucus carota	Carrot	Unassigned	Food Plant	Daucus cyclophilin	171	AEY79726.1	373939374	13
Daucus carota	Carrot	Dau c 1.0101	Food Plant	Daucus Dau c 1	168	AAB01092.1	1335877	7
Daucus carota	Carrot	Dau c 1.0102	Food Plant	Daucus Dau c 1	154	BAA13604.1	1663522	7
Daucus carota	Carrot	Dau c 1.0103	Food Plant	Daucus Dau c 1	154	CAB03715.1	2154732	7
Daucus carota	Carrot	Dau c 1.0104	Food Plant	Daucus Dau c 1	154	CAB03716.1	2154734	7
Daucus carota	Carrot	Dau c 1.0105	Food Plant	Daucus Dau c 1	154	CAB06416.1	2154736	7
Daucus carota	Carrot	Dau c 1.0201	Food Plant	Daucus Dau c 1	154	AAL76932.1	18652047	7
Daucus carota	Carrot	Unassigned	Food Plant	Daucus Dau c 1	154	BAB88129.1	19912791	7
Daucus carota	Carrot	Dau c 1.0301	Food Plant	Daucus Dau c 1	154	ADL32660.1	302379147	12
Daucus carota	Carrot	Unassigned	Food Plant	Daucus Dau c 1	154	ADL32661.1	302379149	12
Daucus carota	Carrot	Unassigned	Food Plant	Daucus Dau c 1	154	ADL32662.1	302379151	12
Daucus carota	Carrot	Unassigned	Food Plant	Daucus Dau c 1	154	ADL32663.1	302379153	12
Daucus carota	Carrot	Unassigned	Food Plant	Daucus Dau c 1	154	ADL32664.1	302379155	12
Daucus carota	Carrot	Unassigned	Food Plant	Daucus Dau c 1	154	ADL32665.1	302379157	12
Daucus carota	Carrot	Unassigned	Food Plant	Daucus Dau c 1	154	ADL32666.1	302379159	12
Daucus carota	Carrot	Dau c 4	Food Plant	Daucus Dau c 4	134	AAL76933.1	18652049	7
Daucus carota	Carrot	Dau c 5.0101	Food Plant	Daucus Dau c 5 isoflavone reductase	306	AEY79728.1	373939378	13
Daucus carota	Carrot	Unassigned	Food Plant	Daucus Dau c 5 isoflavone reductase	306	AEY79727.1	373939376	13
Davidiella tassiana	Fungus	Cla h 10	Aero Fungi	Cladosporium / Davidiella Cla h 10	496	CAA55072.2	76666769	7
Davidiella tassiana	Fungus	Cla h 5.0101	Aero Fungi	Cladosporium / Davidiella Cla h 5	111	CAA55067.2	5777795	10
Davidiella tassiana	Fungus	Cla h 6	Aero Fungi	Cladosporium / Davidiella Cla h 6	440	CAA55070.1	467660	7
Davidiella tassiana	Fungus	Cla h 6	Aero Fungi	Cladosporium / Davidiella Cla h 6	440	P42040.2	6015094	7
Davidiella tassiana	Fungus	Cla h 7.0101	Aero Fungi	Cladosporium / Davidiella Cla h 7	204	CAA55068.1	467629	10
Davidiella tassiana	Fungus	Cla h 8.0101	Aero Fungi	Cladosporium / Davidiella Cla h 8	267	AAO91801.1	37780015	8

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Davidiella tassiana	Fungus	Cla h 9.0101	Aero Fungi	Cladosporium / Davidiella Cla h 9 vacuolar serine	518	AAX14379.1	60116876	10
Davidiella tassiana	Fungus	Unassigned	Aero Fungi	Cladosporium / Davidiella Heat shock 70 kDa protei	643	P40918.1	729764	7
Davidiella tassiana	Fungus	Unassigned	Aero Fungi	Cladosporium / Davidiella Hydrophobin	105	CAD42710.1	22796153	7
Davidiella tassiana	Fungus	Unassigned	Aero Fungi	Cladosporium / Davidiella putative hydrolase	274	ABA42918.1	76446100	10
Davidiella tassiana	Fungus	Unassigned	Aero Fungi	Cladosporium / Davidiella Putative nuclear transpo	125	CAD38166.1	21748151	7
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Der f 36 from Proteome	229	ATI08931.1	1250175279	18
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Der f alpha actinin	885	L7UZ85.1	1160577980	18
Dermatophagoides farinae	House dust mite	Der f 13.0101	Aero Mite	Dermatophagoides Der f 13	131	AAP35078.1	37958167	11
Dermatophagoides farinae	House dust mite	Der f 15	Aero Mite	Dermatophagoides Der f 15 Der p 15	555	AAD52672.1	5815436	7
Dermatophagoides farinae	House dust mite	Der f 16	Aero Mite	Dermatophagoides Der f 16	480	AAM64112.1	21591547	7
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der f 20 / Der p 20	356	AAP57094.1	37785884	8
Dermatophagoides farinae	House dust mite	Der f 20.0201	Aero Mite	Dermatophagoides Der f 20 / Der p 20	356	ABU97470.1	156938897	9
Dermatophagoides farinae	House dust mite	Der f 20.0101	Aero Mite	Dermatophagoides Der f 20 / Der p 20	356	AIO08850.1	685432792	15
Dermatophagoides farinae	House dust mite	Der f 24.0101	Aero Mite	Dermatophagoides Der f 24 and Der p 24 Ubiquinol	118	AGI78542.1	477541860	14
Dermatophagoides farinae	House dust mite	Der f 25.0101	Aero Mite	Dermatophagoides Der f 25	247	AGC56216.1	442565872	14
Dermatophagoides farinae	House dust mite	Der f 25.0201	Aero Mite	Dermatophagoides Der f 25	247	AIO08860.1	685432812	15
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der f 27	427	AAP35082.1	37958175	8
Dermatophagoides farinae	House dust mite	Der f 27.0101	Aero Mite	Dermatophagoides Der f 27	427	AIO08851.1	685432794	15
Dermatophagoides farinae	House dust mite	Der f 28.0101	Aero Mite	Dermatophagoides Der f 28	659	AGC56218.1	442565876	14

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dermatophagoides farinae	House dust mite	Der f 28.0201	Aero Mite	Dermatophagoides Der f 28	654	AIO08848.1	685432788	15
Dermatophagoides farinae	House dust mite	Der f 29.0101	Aero Mite	Dermatophagoides Der f 29	164	AAP35065.1	37958141	8
Dermatophagoides farinae	House dust mite	Der f 30.0101	Aero Mite	Dermatophagoides Der f 30	171	AGC56219.1	442565878	14
Dermatophagoides farinae	House dust mite	Der f 31.0101	Unassigned	Dermatophagoides Der f 31	148	AIO08870.1	685432832	15
Dermatophagoides farinae	House dust mite	Der f 33.0101	Unassigned	Dermatophagoides Der f 33	461	AIO08861.1	685432814	15
Dermatophagoides farinae	House dust mite	Der f 35.0101	Aero Mite	Dermatophagoides Der f 35	143	BAX34757.1	1187443130	18
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der f 5-like	132	BAE45865.1	76880188	7
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides Der f 5-like	132	AAP35068.1	37958147	8
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides Der f 5-like	132	ABO84970.1	140089345	9
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides Der f 5-like	132	ABO84971.1	140089347	9
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides Der f 5-like	132	ABO84972.1	140089349	9
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides Der f 5-like	132	ABO84973.1	140089351	9
Dermatophagoides farinae	House dust mite	Der f 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	321	P16311.2	730035	7
Dermatophagoides farinae	House dust mite	Der f 1.0101	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	321	BAC53948.1	27530349	7
Dermatophagoides farinae	House dust mite	Der f 1.0102	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	276	ABA39436.1	76097507	7
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	321	ABU49605.1	156106765	9
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	263	AAP35075.1	37958161	12
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	305	AFJ68066.1	387178006	13
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	303	ADM52184.1	305387429	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dermatophagoides farinae	House dust mite	Der f 1.0108	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	321	ABL84749.1	119633260	15
Dermatophagoides farinae	House dust mite	Der f 1.0109	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	321	ABL84750.1	119633262	15
Dermatophagoides farinae	House dust mite	Der f 1.0110	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	321	ABL84751.1	119633264	15
Dermatophagoides farinae	House dust mite	Der f 10.0101	Aero Mite	Dermatophagoides Der p 10 / Der f 10	299	BAA04557.1	1359436	7
Dermatophagoides farinae	House dust mite	Der f 11	Aero Mite	Dermatophagoides Der p 11 / Der f 11	692	AAK39511.1	13785807	7
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 11 / Der f 11	876	AIO08864.1	685432820	16
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 14 / Der f 14	341	P39673.1	729979	7
Dermatophagoides farinae	House dust mite	Der f 14.0101	Aero Mite	Dermatophagoides Der p 14 / Der f 14	349	BAA04558.1	1545803	7
Dermatophagoides farinae	House dust mite	Der f 2.0102	Aero Mite	Dermatophagoides Der p 2 / Der f 2	146	BAA01240.1	217306	7
Dermatophagoides farinae	House dust mite	Der f 2.0103	Aero Mite	Dermatophagoides Der p 2 / Der f 2	138	BAA01241.1	217308	7
Dermatophagoides farinae	House dust mite	Der f 2.0105	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	AAL47677.1	17978844	7
Dermatophagoides farinae	House dust mite	Der f 2.0108	Aero Mite	Dermatophagoides Der p 2 / Der f 2	146	CAI05850.1	55859470	7
Dermatophagoides farinae	House dust mite	Der f 2.0107	Aero Mite	Dermatophagoides Der p 2 / Der f 2	146	CAI05849.1	55859468	7
Dermatophagoides farinae	House dust mite	Der f 2.0106	Aero Mite	Dermatophagoides Der p 2 / Der f 2	146	CAI05848.1	55859466	7
Dermatophagoides farinae	House dust mite	Der f 2.0109	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	ABA39438.1	76097511	7
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 2 / Der f 2	146	BAD74060.2	256631558	11
Dermatophagoides farinae	House dust mite	Der f 2.0112	Aero Mite	Dermatophagoides Der p 2 / Der f 2	140	AAP35073.1	37958157	12
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	AFJ68072.1	387178018	13
Dermatophagoides farinae	House dust mite	Der f 2.0101	Aero Mite	Dermatophagoides Der p 2 / Der f 2	138	BAA01239.1	217304	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dermatophagoides farinae	House dust mite	Der f 2.0116	Aero Mite	Dermatophagoides Der p 2 / Der f 2	146	ABN14313.1	124696217	15
Dermatophagoides farinae	House dust mite	Der f 3	Aero Mite	Dermatophagoides Der p 3 / Der f 3	232	AAA99805.1	1314736	7
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 3 / Der f 3	259	ABY28115.1	163638970	9
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 3 / Der f 3	259	ACK76291.1	218203816	10
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 3 / Der f 3	259	ACK76292.1	218203818	10
Dermatophagoides farinae	House dust mite	Der f 3.0101	Aero Mite	Dermatophagoides Der p 3 / Der f 3	259	BAA09920.1	1311457	15
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 6 / Der f 6	20	AAB27594.1	404371	7
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 6 / Der f 6	279	ACK76296.1	218203826	10
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 6 / Der f 6	279	ACK76297.1	218203828	10
Dermatophagoides farinae	House dust mite	Der f 6.0101	Aero Mite	Dermatophagoides Der p 6 / Der f 6	279	AAF28423.1	6808530	11
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 7 / Der f 7	213	AAP35077.1	37958165	8
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 7 / Der f 7	213	ACK76299.1	218203832	10
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 7 / Der f 7	213	AIO08853.1	685432798	16
Dermatophagoides farinae	House dust mite	Der f 18	Aero Mite	Dermatophagoides farinae Der f 18 Der p 18	462	AAM19082.1	27550039	7
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides farinae Der f 21 Chew	136	ABO84963.1	140089314	9
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides farinae Der f 21 Chew	136	ABO84964.1	140089316	9
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides farinae Der f 21 Chew	136	ABO84966.1	140089320	9
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides farinae Der f 21 Chew	136	ABO84967.1	140089322	9
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides farinae Der f 21 Chew	136	ABO84968.1	140089324	9

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides farinae Der f 21 Chew	136	ABO84969.1	140089326	9
Dermatophagoides farinae	House dust mite	Der f 21.0101	Aero Mite	Dermatophagoides farinae Der f 21 Chew	136	AHC94806.1	567768173	15
Dermatophagoides farinae	House dust mite	Der f 34.0101	Unassigned	Dermatophagoides farinae Der f 34	128	BAV90601.1	1098871171	17
Dermatophagoides farinae	House dust mite	Der f 4.0101	Unassigned	Dermatophagoides farinae Der f 4	525	AHX03180.1	612487835	15
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Der f 4	525	AIP86946.1	685848330	16
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Der f 4	525	AIP86945.1	685848328	16
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Der f 4	525	AIP86944.1	685848326	16
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Der f 4	525	AIP86943.1	685848324	16
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Der f 4	525	AIP86942.1	685848322	16
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Der f 4	525	AIP86941.1	685848320	16
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Der f 4	525	AIP86940.1	685848318	16
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Der f 4	525	AIP86939.1	685848316	16
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Der f 4	525	AJF93907.1	751425403	16
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Pseudo-Der f 8	219	AAP35080.1	37958171	12
Dermatophagoides farinae	House dust mite	Unassigned	Unassigned	Dermatophagoides farinae Pseudo-Der f 8	221	AIO08867.1	685432826	16
Dermatophagoides farinae	House dust mite	Unassigned	Aero Mite	Dermatophagoides Profilin	130	AIO08866.1	685432824	16
Dermatophagoides microceras	House dust mite	Der m 1.0101	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	30	P16312.1	127205	7
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Unassigned	Der p 36	227	ATI08932.1	1250175281	18

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dermatophagoides pteronyssinus	House dust mite	Der p 15.0101	Aero Mite	Dermatophagoides Der f 15 Der p 15	532	AAY84565.1	67975089	7
Dermatophagoides pteronyssinus	House dust mite	Der p 15.0102	Aero Mite	Dermatophagoides Der f 15 Der p 15	558	AAY84564.2	78128018	7
Dermatophagoides pteronyssinus	House dust mite	Der p 20.0101	Aero Mite	Dermatophagoides Der f 20 / Der p 20	356	ACD50950.1	188485735	10
Dermatophagoides pteronyssinus	House dust mite	Der p 24.0101	Aero Mite	Dermatophagoides Der f 24 and Der p 24 Ubiquinol	118	ALA65345.1	922664427	16
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der f 30	180	AAG02250.1	15072346	7
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38361.1	21725560	7
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38362.1	21725562	7
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38363.1	21725564	7
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38364.1	21725566	7
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38365.1	21725568	7
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38366.1	21725570	7
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38367.1	21725572	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38368.1	21725574	7
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38369.1	21725576	7
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38370.1	21725578	7
Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	CAD38371.1	21725580	7
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	216	AAX47076.1	61608445	7
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	2AS8_B	83754033	7
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	223	ABV66255.1	157696052	9
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	222	3F5V_B	223365887	10
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	320	ACG58378.1	195933901	10
Dermatophagoides pteronyssinus	House dust mite	Der p 1.0124	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	302	CAQ68250.1	256095986	11
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	96	AAA28296.1	387592	11
Dermatophagoides pteronyssinus	House dust mite	Der p 1.0101	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	320	AAB60215.1	511953	12

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	304	AFJ68065.1	387178004	13
Dermatophagoides pteronyssinus	House dust mite	Der p 1.0113	Aero Mite	Dermatophagoides Der p 1 Der f 1 Der m 1	302	ABA39435.1	76097505	15
Dermatophagoides pteronyssinus	House dust mite	Der p 10	Aero Mite	Dermatophagoides Der p 10 / Der f 10	284	AAB69424.1	2353266	7
Dermatophagoides pteronyssinus	House dust mite	Der p 10.0101	Aero Mite	Dermatophagoides Der p 10 / Der f 10	284	CAA75141.1	2440053	7
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 10 / Der f 10	281	ABB52642.1	80553470	7
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 10 / Der f 10	284	ACI32128.1	208970286	10
Dermatophagoides pteronyssinus	House dust mite	Der p 11.0101	Aero Mite	Dermatophagoides Der p 11 / Der f 11	875	AAO73464.1	37778944	7
Dermatophagoides pteronyssinus	House dust mite	Der p 13.0101	Aero Mite	Dermatophagoides Der p 13	131	ADK92390.1	302035350	12
Dermatophagoides pteronyssinus	House dust mite	Der p 14.0101	Aero Mite	Dermatophagoides Der p 14 / Der f 14	1662	AAM21322.1	20385544	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	1KTJ_A	21465915	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38372.1	21725582	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38373.1	21725584	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38374.1	21725586	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38375.1	21725588	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38376.1	21725590	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38377.1	21725592	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38378.1	21725594	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38379.1	21725596	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38381.1	21725600	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38382.1	21725602	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAD38383.1	21725604	7
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	ABA39437.1	76097509	7
Dermatophagoides pteronyssinus	House dust mite	Der p 2.0114	Aero Mite	Dermatophagoides Der p 2 / Der f 2	146	CAK22338.1	99644635	7
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 2 / Der f 2	130	ABG76196.1	110560872	9

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	1A9V_A	157829757	9
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 2 / Der f 2	145	ABY53034.1	164415595	9
Dermatophagoides pteronyssinus	House dust mite	Der p 2.0101	Aero Mite	Dermatophagoides Der p 2 / Der f 2	145	AAF86462.1	9280543	10
Dermatophagoides pteronyssinus	House dust mite	Der p 2.0110	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	CAQ68249.1	256095984	11
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	AFJ68070.1	387178014	13
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 2 / Der f 2	129	AFJ68067.1	387178008	13
Dermatophagoides pteronyssinus	House dust mite	Der p 21.0101	Aero Mite	Dermatophagoides Der p 21	140	ABC73706.1	85687540	7
Dermatophagoides pteronyssinus	House dust mite	Der p 23.0101	Aero Mite	Dermatophagoides Der p 23 Peritrophin-like protein	90	ACB46292.1	171466145	14
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Unassigned	Dermatophagoides Der p 23 Peritrophin-like protein	50	4ZCE_A	955264737	17
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Unassigned	Dermatophagoides Der p 23 Peritrophin-like protein	99	ALA22869.1	920684621	17
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Unassigned	Dermatophagoides Der p 23 Peritrophin-like protein	98	ALA22868.1	920684619	17
Dermatophagoides pteronyssinus	House dust mite	Der p 3	Aero Mite	Dermatophagoides Der p 3 / Der f 3	261	AAA19973.1	511476	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dermatophagoides pteronyssinus	House dust mite	Der p 4	Aero Mite	Dermatophagoides Der p 4	496	AAD38942.1	5059162	7
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 4	19	P49274.1	1351935	7
Dermatophagoides pteronyssinus	House dust mite	Der p 5.0102	Aero Mite	Dermatophagoides Der p 5	132	AAB32842.1	913285	7
Dermatophagoides pteronyssinus	House dust mite	Der p 5	Aero Mite	Dermatophagoides Der p 5	132	CAD69036.1	28798085	7
Dermatophagoides pteronyssinus	House dust mite	Der p 5.0101	Aero Mite	Dermatophagoides Der p 5	148	CAA35692.1	9072	15
Dermatophagoides pteronyssinus	House dust mite	Der p 6	Aero Mite	Dermatophagoides Der p 6 / Der f 6	20	P49277.1	1352239	7
Dermatophagoides pteronyssinus	House dust mite	Der p 7.0101	Aero Mite	Dermatophagoides Der p 7 / Der f 7	215	AAA80264.1	1045602	7
Dermatophagoides pteronyssinus	House dust mite	Der p 7	Aero Mite	Dermatophagoides Der p 7 / Der f 7	215	CAC09234.1	10189811	7
Dermatophagoides pteronyssinus	House dust mite	Der f 7.0101	Aero Mite	Dermatophagoides Der p 7 / Der f 7	213	AAB35977.1	1311689	10
Dermatophagoides pteronyssinus	House dust mite	Der p 8.0101	Aero Mite	Dermatophagoides Der p 8	219	AAB32224.1	807138	7
Dermatophagoides pteronyssinus	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 8	219	AAX37326.1	60920878	7
Dermatophagoides pteronyssinus	House dust mite	Der p 18.0101	Aero Mite	Dermatophagoides farinae Der f 18 Der p 18	462	AAY84563.1	67975085	7
Dermatophagoides siboney	House dust mite	Unassigned	Aero Mite	Dermatophagoides Der p 2 / Der f 2	146	ABC96702.1	86450747	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Dolichovespula arenaria	Yellow jacket	Dol a 5.0101	Venom or Salivary	Dolichovespula Venom allergen 5	203	AAA28303.1	156719	11
Dolichovespula maculata	Whiteface hornet	Dol m 1.02	Venom or Salivary	Dolichovespula Dol m 1 Phospholipase A1B	303	P53357.1	1709542	7
Dolichovespula maculata	Whiteface hornet	Dol m 1.0101	Venom or Salivary	Dolichovespula Dol m 1 Phospholipase A1B	317	CAA47341.1	288917	8
Dolichovespula maculata	Whiteface hornet	Dol m 2.0101	Venom or Salivary	Dolichovespula Dol m 2 Hyaluronidase	331	AAA68279.1	511604	11
Dolichovespula maculata	Whiteface hornet	Dol m 5.0101	Venom or Salivary	Dolichovespula Venom allergen 5	227	AAA28301.1	156715	11
Dolichovespula maculata	Whiteface hornet	Dol m 5.02	Venom or Salivary	Dolichovespula Venom allergen 5	212	AAA28302.1	552080	11
Epicoccum nigrum	Fungus	Epi p 1.0101	Aero Fungi	Epicoccum Epi p 1	18	P83340.1	24636820	9
Equus caballus	Horse	Equ c 1.0101	Aero Animal	Equus Equ c 1	187	AAC48691.1	1575778	11
Equus caballus	Horse	Equ c 2.0101	Aero Animal	Equus Equ c 2	29	P81216.1	3121755	7
Equus caballus	Horse	Equ c 2.0102	Aero Animal	Equus Equ c 2	19	P81217.1	3121756	7
Equus caballus	Horse	Equ c 3.0101	Aero Animal	Equus Equ c 3	607	CAA52194.1	399672	7
Equus caballus	Horse	Equ c 4.0101	Aero Animal	Equus Equ c 4 and Equ c 5	228	AAM09530.3	126514234	8
Erimacrus isenbeckii	Horsehair crab	Unassigned	Food Animal	Erimacrus tropomyosin	284	BAF47268.1	125995169	8
Erimacrus isenbeckii	Horsehair crab	Unassigned	Food Animal	Erimacrus tropomyosin	284	BAF47269.1	125995171	8
Eriocheir sinensis	Chinese mitten crab	Unassigned	Unassigned	Eriocheir sinensis Eri s 2	252	AAO73305.1	37778438	16
Eriocheir sinensis	Chinese mitten crab	Unassigned	Food Animal	Eriocheir tropomyosin	284	ABO71783.1	134305330	8
Euphausia pacifica	North Pacific Krill	Unassigned	Food Animal	Euphausia	284	BAF76431.1	156712754	9
Euphausia superba	Krill	Unassigned	Food Animal	Euphausia	284	BAF76430.1	156712752	9
Euroglyphus maynei	House dust mite	Eur m 1.0101	Aero Mite	Euroglyphus Eur m 1	321	AAC82351.1	3941388	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Euroglyphus maynei	House dust mite	Unassigned	Aero Mite	Euroglyphus Eur m 1	327	AAC82352.1	3941390	7
Euroglyphus maynei	House dust mite	Eur m 2.0102	Aero Mite	Euroglyphus Eur m 2	135	AAC82350.1	3941386	7
Euroglyphus maynei	House dust mite	Eur m 2.0101	Aero Mite	Euroglyphus Eur m 2	145	AAC82349.1	3941384	11
Evynnis japonica	Crimson seabream	Unassigned	Food Animal	Evynnis parvalbumin	109	BAK09233.1	327342663	12
Evynnis japonica	Crimson seabream	Unassigned	Food Animal	Evynnis parvalbumin	108	BAK09232.1	327342661	12
Fagopyrum esculentum	Buckwheat	Unassigned	Food Plant	Fagopyrum BW 8 kDa protein	133	BAB79444.1	17907758	7
Fagopyrum esculentum	Buckwheat	Unassigned	Food Plant	Fagopyrum esculentum 13S globulins IgE binding	453	BAO50872.1	584592120	15
Fagopyrum esculentum	Buckwheat	Unassigned	Food Plant	Fagopyrum esculentum 13S globulins IgE binding	453	BAO50870.1	584592116	15
Fagopyrum esculentum	Buckwheat	Unassigned	Food Plant	Fagopyrum Fag e 2 Fag t 2	127	AAX57578.1	61970231	7
Fagopyrum esculentum	Buckwheat	Fag e 2.0101	Food Plant	Fagopyrum Fag e 2 Fag t 2	149	ABC18306.1	83416591	7
Fagopyrum esculentum	Buckwheat	Unassigned	Food Plant	Fagopyrum Legumin-like protein	565	O23878.1	29839254	9
Fagopyrum esculentum	Buckwheat	Unassigned	Food Plant	Fagopyrum Legumin-like protein	504	O23880.1	29839255	9
Fagopyrum esculentum	Buckwheat	Unassigned	Food Plant	Fagopyrum Legumin-like protein	538	Q9XFM4.1	29839419	9
Fagopyrum esculentum	Buckwheat	Fag e 3.0101	Food Plant	Fagopyrum vicilin-like Fag e 3	136	ABQ10638.1	146217148	9
Fagopyrum esculentum	Buckwheat	Unassigned	Unassigned	Fagopyrum vicilin-like Fag e 3	136	BAT21117.1	939106201	17
Fagopyrum tataricum	Buckwheat	Unassigned	Food Plant	Fagopyrum BW 8 kDa protein	133	ABO93594.1	144228127	8
Fagopyrum tataricum	Buckwheat	Fag t 2.0101	Food Plant	Fagopyrum Fag e 2 Fag t 2	149	ADW27428.1	320445237	12
Fagopyrum tataricum	Buckwheat	Unassigned	Food Plant	Fagopyrum Legumin-like protein	515	ABI32184.1	113200131	9
Fagus sylvatica	European Beech	Unassigned	Aero Plant	Fagus Fag s 1	160	ACJ23865.1	212291472	10

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Fagus sylvatica	European Beech	Fag s 1	Aero Plant	Fagus Fag s 1	160	ACJ23864.1	212291470	10
Fagus sylvatica	European Beech	Unassigned	Aero Plant	Fagus Fag s 1	160	ACJ23866.1	212291474	10
Farfantepenaeus aztecus	Brown shrimp	Pen a 1	Food Animal	Farfantepenaeus Pen a 1	284	AAZ76743.1	73532979	7
Felis catus	Cat	Fel d 1	Aero Animal	Felis Fel d 1 Chain 1	88	CAA44343.1	1364212	7
Felis catus	Cat	Fel d 1	Aero Animal	Felis Fel d 1 Chain 1	92	CAA44344.1	1364213	7
Felis catus	Cat	Fel d 1	Aero Animal	Felis Fel d 1 Chain 1	92	P30438.2	1169665	7
Felis catus	Cat	Fel d 1.0101	Aero Animal	Felis Fel d 1 Chain 1	92	AAC37318.1	163825	7
Felis catus	Cat	Unassigned	Aero Animal	Felis Fel d 1 Chain 1	88	NP_001041618.1	114326420	8
Felis catus	Cat	Unassigned	Aero Animal	Felis Fel d 1 chain 2	107	CAA44345.1	395407	8
Felis catus	Cat	Fel d 1.0101	Aero Animal	Felis Fel d 1 chain 2	109	AAC41616.1	163823	12
Felis catus	Cat	Fel d 2.0101	Aero Animal	Felis Fel d 2	608	CAA59279.1	886485	7
Felis catus	Cat	Fel d 3	Aero Animal	Felis Fel d 3	98	AAL49391.1	17939981	7
Felis catus	Cat	Fel d 4	Aero Animal	Felis Fel d 4	186	AAS77253.1	45775300	7
Felis catus	Cat	Fel d 7.0101	Aero Animal	Felis Fel d 7	180	ADK56160.1	301072397	12
Felis catus	Cat	Fel d 8.0101	Aero Animal	Felis Fel d 8 latherin-like	228	ADM15668.1	303387468	12
Fenneropenaeus chinensis	Chinese white shrimp	Unassigned	Food Animal	Fenneropenaeus Arginine kinase	53	AAS98889.1	46486948	9
Fenneropenaeus chinensis	Chinese white shrimp	Unassigned	Food Animal	Fenneropenaeus Arginine kinase	53	AAS98890.1	46486951	9
Fenneropenaeus merguiensis	Banana Prawn	Unassigned	Food Animal	Fenneropenaeus hemocyanin banana shrimp	661	AGT20779.1	530340505	15
Fenneropenaeus merguiensis	Banana Prawn	Unassigned	Food Animal	Fenneropenaeus enolase	117	AEM89226.1	344049993	15
Forcipomyia taiwana	biting midges	For t 1.0101	Venom or Salivary	Forcipomyia For t 1	118	ACD65080.1	188572341	10
Forcipomyia taiwana	biting midges	For t 1.0101	Venom or Salivary	Forcipomyia For t 2	325	ACD65081.1	188572343	10
Fragaria x ananassa	Strawberry	Fra a 1	Food Plant	Fragaria Fra a 1	160	CAJ85646.1	90185692	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Fragaria x ananassa	Strawberry	Fra a 1	Food Plant	Fragaria Fra a 1	159	CAJ85644.1	90185688	7
Fragaria x ananassa	Strawberry	Fra a 1	Food Plant	Fragaria Fra a 1	160	CAJ85642.1	90185684	7
Fragaria x ananassa	Strawberry	Fra a 1	Food Plant	Fragaria Fra a 1	160	CAJ85641.1	90185682	7
Fragaria x ananassa	Strawberry	Fra a 1	Food Plant	Fragaria Fra a 1	160	ABD39049.1	88082485	7
Fragaria x ananassa	Strawberry	Unassigned	Unassigned	Fragaria Fra a 1	160	ACX47057.1	260600660	11
Fragaria x ananassa	Strawberry	Unassigned	Unassigned	Fragaria Fra a 1	159	ACX47058.1	260600662	11
Fragaria x ananassa	Strawberry	Unassigned	Food Plant	Fragaria Fra a 1	162	4C9C_B	550544407	15
Fragaria x ananassa	Strawberry	Fra a 3.0101	Food Plant	Fragaria Fra a 3	117	CAC86258.1	18477856	15
Fragaria x ananassa	Strawberry	Fra a 3.0102	Food Plant	Fragaria Fra a 3	117	AAY83342.1	67937767	15
Fragaria x ananassa	Strawberry	Fra a 3.0201	Food Plant	Fragaria Fra a 3	117	AAY83341.1	67937765	15
Fragaria x ananassa	Strawberry	Fra a 3.0202	Food Plant	Fragaria Fra a 3	117	AAY83345.1	67937773	15
Fraxinus excelsior	European ash	Unassigned	Aero Plant	Fraxinus excelsior polcalcin	84	AHL24661.1	589912891	15
Fraxinus excelsior	European ash	Unassigned	Aero Plant	Fraxinus excelsior profilin	134	AHL24660.1	589912889	15
Fraxinus excelsior	European ash	Fra e 1.0201	Aero Plant	Fraxinus Fra e 1	146	AAQ83588.1	34978692	7
Fraxinus excelsior	European ash	Fra e 1.0102	Aero Plant	Fraxinus Fra e 1	145	AAV74343.1	56122438	7
Fraxinus excelsior	European ash	Fra e 1.0101	Aero Plant	Fraxinus Fra e 1	145	AAQ08947.1	33327133	7
Fulvia mutica	Mollusc	Unassigned	Food Animal	Fulvia tropomyosin	284	BAH10153.1	219806596	10
Fusarium culmorum	Fungus	Unassigned	Aero Fungi	Fusarium claimed Fus c 3	450	AAN73248.1	25361513	7
Fusarium culmorum	Fungus	Fus c 1	Aero Fungi	Fusarium Fus c 1	109	AAL79930.1	19879657	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Fusarium culmorum	Fungus	Fus c 2	Aero Fungi	Fusarium Fus c 2	121	AAL79931.1	19879659	7
Fusarium proliferatum	Fungus	Fus p 4.0101	Aero Fungi	Fusarium Fus p 4	323	AHY02994.1	619498167	15
Gadus callarias	Baltic cod	Gad c 1	Food Animal	Gadus Gad c 1 Gad m 1	113	P02622.1	131112	7
Gadus morhua	Atlantic cod	Gad m 1.0101	Food Animal	Gadus Gad c 1 Gad m 1	109	AAK63086.1	14531014	7
Gadus morhua	Atlantic cod	Gad m 1.0201	Food Animal	Gadus Gad c 1 Gad m 1	109	AAK63087.1	14531016	7
Gadus morhua	Atlantic cod	Gad m 1.0102	Food Animal	Gadus Gad c 1 Gad m 1	109	CAM56785.1	148356691	9
Gadus morhua	Atlantic cod	Gad m 1.0202	Food Animal	Gadus Gad c 1 Gad m 1	109	CAM56786.1	148356693	9
Gadus morhua	Atlantic cod	Gad m 2.0101	Food Animal	Gadus Morhua Gad m 2	11	B3A0L6.1	576011130	15
Gadus morhua	Atlantic cod	Gad m 3.0101	Food Animal	Gadus morhua Gad m 3	15	P86980.1	576011086	15
Gallus gallus	Chicken	Gal d 9	Unassigned	Gallus enolase Gal d 9	434	NP_990450.1	46048765	18
Gallus gallus	Chicken	Gal d 1	Food Animal	Gallus Gal d 1	210	P01005.1	124757	7
Gallus gallus	Chicken	Unassigned	Food Animal	Gallus Gal d 1	210	ACJ04729.1	209979542	10
Gallus gallus	Chicken	Gal d 2	Food Animal	Gallus Gal d 2	155	CAA23681.1	63052	7
Gallus gallus	Chicken	Gal d 2.0101	Food Animal	Gallus Gal d 2	386	P01012.2	129293	7
Gallus gallus	Chicken	Gal d 2	Food Animal	Gallus Gal d 2	386	CAA23682.1	808969	7
Gallus gallus	Chicken	Gal d 2	Food Animal	Gallus Gal d 2	385	1JTI_A	15826578	7
Gallus gallus	Chicken	Unassigned	Food Animal	Gallus Gal d 2	385	1UHG_D	34811333	7
Gallus gallus	Chicken	Gal d 3.0101	Food Animal	Gallus Gal d 3	705	CAA26040.1	757851	7
Gallus gallus	Chicken	Gal d 3	Food Animal	Gallus Gal d 3	705	P02789.2	1351295	7
Gallus gallus	Chicken	Gal d 4	Food Animal	Gallus Gal d 4	147	P00698.1	126608	7
Gallus gallus	Chicken	Gal d 4	Food Animal	Gallus Gal d 4	24	AAA48944.1	212279	7
Gallus gallus	Chicken	Gal d 4.0101	Food Animal	Gallus Gal d 4	147	CAA23711.1	63581	15
Gallus gallus	Chicken	Gal d 5	Food Animal	Gallus Gal d 5	615	CAA43098.1	63748	7
Gallus gallus	Chicken	Unassigned	Food Animal	Gallus Gal d 6 YGP42	284	BAA13973.1	3	14
Gallus gallus	Chicken	Gal d 9	Food Animal	Gallus gallus Gal d 7	192	P02604.3	55584149	16
Gallus gallus	Chicken	Gal d 8.0101	Food Animal	Gallus parvalbumin Gal d 8	110	CAX32963.1	225877920	10

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Glossina morsitans morsitans	Tsetse fly	Unassigned	Venom or Salivary	Glossina Glo m 5	258	ADD18879.1	289740263	11
Glossina morsitans morsitans	Tsetse fly	Unassigned	Venom or Salivary	Glossina Glo m 5	259	ADD19985.1	289742475	11
Glossina morsitans morsitans	Tsetse fly	Unassigned	Venom or Salivary	Glossina Glo m 5	222	ADD19989.1	289742483	11
Glossina morsitans morsitans	Tsetse fly	Glo m 5.0101	Venom or Salivary	Glossina Glo m 5	259	AAF82096.1	8927462	11
Glycine max	Soybean	Gly m 7.0101	Unassigned	Glycine 68kDa biotinylated protein	643	ACS49840.1	240254706	11
Glycine max	Soybean	Gly m 1.0101	Food Plant	Glycine Gly m 1	80	P24337.1	123506	12
Glycine max	Soybean	Gly m 3.0102	Food Plant	Glycine Gly m 3	131	CAA11755.1	3021373	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 3	131	ABU97472.1	156938901	9
Glycine max	Soybean	Gly m 3.0101	Food Plant	Glycine Gly m 3	131	CAA11756.1	3021375	15
Glycine max	Soybean	Gly m 4	Food Plant	Glycine Gly m 4	158	CAA42646.1	18744	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 5.0101 alpha subunit beta congl	605	CAA35691.1	18536	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 5.0101 alpha subunit beta congl	218	AAA33947.1	169927	7
Glycine max	Soybean	Gly m 5.0101	Food Plant	Glycine Gly m 5.0101 alpha subunit beta congl	543	BAA23360.2	9967357	15
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 5.0201 alpha prime beta congly	639	AAB01374.1	169929	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 5.0201 alpha prime beta congly	621	BAB64303.1	15425631	15
Glycine max	Soybean	Gly m 5.0201	Food Plant	Glycine Gly m 5.0201 alpha prime beta congly	559	BAA74452.2	9967361	15
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 5.0301 beta sub unit beta congl	439	BAB64306.1	15425637	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Glycine max	Soybean	Gly m 5.0301	Food Plant	Glycine Gly m 5.0301 beta sub unit beta congl	439	P25974.1	121282	15
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 6.0101	495	CAA26723.1	18615	7
Glycine max	Soybean	Gly m 6.0101	Food Plant	Glycine Gly m 6.0101	495	AAA33966.1	169973	15
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 6.0201	485	CAA26575.1	18609	7
Glycine max	Soybean	Gly m 6.0201	Food Plant	Glycine Gly m 6.0201	485	BAA00154.1	218265	15
Glycine max	Soybean	Gly m 6.0301	Food Plant	Glycine Gly m 6.0301	481	CAA33217.1	18639	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 6.0401	562	CAA37044.1	18641	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 6.0401	562	CAA26478.1	732706	7
Glycine max	Soybean	Gly m 6.0401	Food Plant	Glycine Gly m 6.0401	563	BAA74953.1	4249568	15
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 6.0501	516	AAA33964.1	169969	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 6.0501	240	AAA33965.1	169971	7
Glycine max	Soybean	Gly m 6.0501	Food Plant	Glycine Gly m 6.0501	517	BAB15802.1	10566449	15
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m 8 2s albumin	155	AAD09630.1	4097894	14
Glycine max	Soybean	Gly m 8.0101	Food Plant	Glycine Gly m 8 2s albumin	158	NP_001238443.1	351727517	15
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m Bd 28K	373	ACD36976.1	187766751	10
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m Bd 28K	373	ACD36975.1	187766749	10
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m Bd 28K	373	ACD36974.1	187766747	10
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m Bd 28K	455	ACD36978.1	187766755	10
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m Bd 28K	476	BAB21619.2	410067729	15
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m Bd 30 kDa	379	P22895.1	129353	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m Bd 30 kDa	379	AAB09252.1	1199563	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Gly m Bd 30 kDa	379	BAA25899.1	3097321	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Major Gly 50 kDa allergen	17	P82947.1	85681057	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Trypsin inhibitor	217	CAA45777.1	18770	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Trypsin inhibitor	217	CAA45778.1	18772	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Trypsin inhibitor	216	AAB23464.1	256429	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Trypsin inhibitor	203	AAB23482.1	256635	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Glycine max	Soybean	Unassigned	Food Plant	Glycine Trypsin inhibitor	204	AAB23483.1	256636	7
Glycine max	Soybean	Unassigned	Food Plant	Glycine Trypsin inhibitor	208	CAA56343.1	510515	7
Glycine soja	Soybean	Unassigned	Food Plant	Glycine Gly m 6.0401	563	CAA60533.1	806556	7
Glycyphagus domesticus	Storage mite	Gly d 2.0101	Aero Mite	Glycyphagus Gly d 2	128	CAB59976.1	6179520	7
Glycyphagus domesticus	Storage mite	Gly d 2.0201	Aero Mite	Glycyphagus Gly d 2	125	CAB76459.1	7160811	7
Glycyphagus domesticus	Storage mite	Unassigned	Aero Mite	Glycyphagus Gly d 2	141	AAQ54603.1	33772588	7
Haliotis discus discus	Disk abalone	Unassigned	Food Animal	Haliotis Hal m 1 tropomyosin	284	BAH10148.1	219806586	10
Haliotis discus discus	Disk abalone	Unassigned	Food Animal	Haliotis paramyosin	860	BAJ61596.1	318609972	12
Haliotis diversicolor	Abalone	Unassigned	Food Animal	Haliotis Hal m 1 tropomyosin	284	AAG08987.1	9954249	7
Haliotis laevigata x Haliotis rubra	Jade Abalone	Hal l 1	Food Animal	Haliotis Hal m 1 tropomyosin	284	APG42675.1	1109557549	18
Helianthus annuus	Sunflower	Hel a 2	Aero Plant	Helianthus Hel a 2	133	CAA75506.1	3581965	7
Helianthus annuus	Sunflower	Hel a 3.0101	Food Plant	Helianthus Hel a 3	116	AAP47226.1	31324341	15
Helianthus annuus	Sunflower	Unassigned	Food Plant	Helianthus Seed 2S albumin	141	P23110.1	112745	9
Helix aspersa	Brown garden snail	Hel as 1.0101	Food Animal	Helix Hel as 1 tropomyosin	284	CAB38044.1	4468224	7
Hevea brasiliensis	Para rubber tree	Hev b 1.0101	Contact	Hevea Hev b 1	138	CAA39880.1	18839	15
Hevea brasiliensis	Para rubber tree	Hev b 10.0101	Contact	Hevea Hev b 10	233	AAA16792.1	348137	7
Hevea brasiliensis	Para rubber tree	Hev b 10.0102	Contact	Hevea Hev b 10	205	CAB53458.1	5777414	7
Hevea brasiliensis	Para rubber tree	Hev b 10.0103	Contact	Hevea Hev b 10	205	CAC13961.1	10862818	7
Hevea brasiliensis	Para rubber tree	Hev b 11.0101	Contact	Hevea Hev b 11	295	CAC42881.1	14575525	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Hevea brasiliensis	Para rubber tree	Hev b 12	Contact	Hevea Hev b 12	116	AAL25839.1	20135538	7
Hevea brasiliensis	Para rubber tree	Hev b 13	Contact	Hevea Hev b 13	391	AAP37470.1	30909057	7
Hevea brasiliensis	Para rubber tree	Hev b 14.0101	Contact	Hevea Hev b 14 hevamine	208	ADR82196.1	313870530	12
Hevea brasiliensis	Para rubber tree	Hev b 15.0101	Contact	Hevea Hev b 15	70	CCW27997.1	571257122	15
Hevea brasiliensis	Para rubber tree	Hev b 2.0101	Contact	Hevea Hev b 2	374	AAA87456.1	1184668	7
Hevea brasiliensis	Para rubber tree	Hev b 2	Contact	Hevea Hev b 2	374	AAP87281.1	32765543	7
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 2	374	ABN03965.1	124294783	8
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 2	374	ABN03966.1	124294785	8
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 2	374	ABN09653.1	124365249	8
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 2	374	ABN09654.1	124365251	8
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 2	374	ABN09655.1	124365253	8
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 2	374	ACY91851.1	268037674	11
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 2	374	ACZ74626.1	270315180	11
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 2	373	AEV41413.1	359359690	13
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 2	374	AFJ97275.1	387778882	13
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 2	374	AFJ97274.1	387778880	13
Hevea brasiliensis	Para rubber tree	Hev b 3.0101	Contact	Hevea Hev b 3	204	AAC82355.1	3818475	11
Hevea brasiliensis	Para rubber tree	Hev b 4.0101	Contact	Hevea Hev b 4	366	AAR98518.1	46410859	7
Hevea brasiliensis	Para rubber tree	Hev b 5	Contact	Hevea Hev b 5	151	AAC49447.1	1480457	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Hevea brasiliensis	Para rubber tree	Hev b 6	Contact	Hevea Hev b 6	187	CAA05978.1	2832430	7
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 6	43	1WKX_A	73535415	7
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 6	204	ABW34946.1	158342650	9
Hevea brasiliensis	Para rubber tree	Hev b 7.01	Contact	Hevea Hev b 7	388	AAC27724.1	1916805	7
Hevea brasiliensis	Para rubber tree	Hev b 7.02	Contact	Hevea Hev b 7	388	CAA11041.1	3087805	7
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 7	388	CAA11042.1	3288200	7
Hevea brasiliensis	Para rubber tree	Hev b 7	Contact	Hevea Hev b 7	388	AAF25553.1	6707018	7
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 7	387	CAE85467.1	41581137	7
Hevea brasiliensis	Para rubber tree	Hev b 8.0101	Contact	Hevea Hev b 8	131	CAA75312.1	3183706	7
Hevea brasiliensis	Para rubber tree	Hev b 8	Contact	Hevea Hev b 8	131	1G5U_A	11513601	7
Hevea brasiliensis	Para rubber tree	Hev b 8.0201	Contact	Hevea Hev b 8	131	AAF34341.1	6979167	11
Hevea brasiliensis	Para rubber tree	Hev b 8.0202	Contact	Hevea Hev b 8	131	AAF34342.1	6979169	11
Hevea brasiliensis	Para rubber tree	Hev b 8.0203	Aero Mite	Hevea Hev b 8	131	AAF34343.1	6979171	11
Hevea brasiliensis	Para rubber tree	Hev b 8.0102	Contact	Hevea Hev b 8	131	CAB51914.1	5689740	15
Hevea brasiliensis	Para rubber tree	Hev b 8.0204	Contact	Hevea Hev b 8	131	CAB96215.1	8919948	15
Hevea brasiliensis	Para rubber tree	Hev b 9	Contact	Hevea Hev b 9	445	CAC00532.1	9581744	7
Hevea brasiliensis	Para rubber tree	Unassigned	Contact	Hevea Hev b 9	445	Q9LEI9.1	14423687	9
Hevea brasiliensis subsp. brasiliensis	Para rubber tree	Hev b 11.0102	Contact	Hevea Hev b 11	295	CAD24068.1	27526732	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Holcus lanatus	Velvet grass	Hol l 1.0101	Aero Plant	Holcus Hol l 1	265	CAA81610.1	414703	7
Holcus lanatus	Velvet grass	Hol l 1.0102	Aero Plant	Holcus Hol l 1	248	CAA93121.1	1167836	7
Holcus lanatus	Velvet grass	Unassigned	Aero Plant	Holcus Hol l 1	263	CAA10140.1	3860384	7
Holcus lanatus	Velvet grass	Unassigned	Aero Plant	Holcus Hol l 5	20	Q7M262	75140046	7
Holcus lanatus	Velvet grass	Hol l 5.0201	Aero Plant	Holcus Hol l 5	240	CAB10766.1	2266623	7
Holcus lanatus	Velvet grass	Hol l 5.0101	Aero Plant	Holcus Hol l 5	264	CAB10765.1	2266625	7
Holcus lanatus	Velvet grass	Unassigned	Aero Plant	Holcus Hol l 5	296	AAG42255.1	11991229	7
Homarus americanus	American lobster	Hom a 1.0102	Food Animal	Homarus Hom a 1	284	AAC48288.1	2660868	7
Homarus americanus	American lobster	Hom a 1.0101	Food Animal	Homarus Hom a 1	284	AAC48287.1	2660866	15
Hordeum vulgare	Barley	Unassigned	Aero Plant	Hordeum Alpha-amylase inhibitor component CMb	149	P32936.2	585290	7
Hordeum vulgare	Barley	Unassigned	Food Plant	Hordeum Hor v 20	289	P80198.1	1708280	15
Hordeum vulgare	Barley	Hor v 20.0101	Food Plant	Hordeum Hor v 20	286	CAA51204.1	288709	15
Hordeum vulgare	Barley	Unassigned	Food Plant	Hordeum LTP 1	134	CAA42832.1	19039	7
Hordeum vulgare	Barley	Unassigned	Aero Plant	Hordeum LTP 1	117	AAA32970.1	167077	7
Hordeum vulgare	Barley	Unassigned	Aero Plant	Hordeum Trypsin inhibitor CMe	144	CAA35188.1	1405736	7
Hordeum vulgare subsp. vulgare	Barley	Unassigned	Aero Plant	Hordeum Alpha-amylase inhibitor BDAI-1	152	CAA08836.1	3367714	7
Hordeum vulgare subsp. vulgare	Barley	Unassigned	Aero Plant	Hordeum Alpha-amylase inhibitor component Cma	144	CAA41956.1	18955	7
Hordeum vulgare subsp. vulgare	Barley	Unassigned	Aero Plant	Hordeum Alpha-amylase inhibitor component Cma	145	CAA49555.1	439275	7
Hordeum vulgare subsp. vulgare	Barley	Hor v 15.0101	Food Plant	Hordeum Hor v 15	146	CAA45085.1	19003	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Hordeum vulgare subsp. vulgare	Barley	Unassigned	Aero Plant	Hordeum Trypsin inhibitor CMe	148	CAA46705.1	19009	7
Humulus japonicus	Japanese hop	Hum j 1	Aero Plant	Humulus Humj1	155	AAP94213.1	33113263	7
Humulus scandens	Japanese hop	Unassigned	Aero Plant	Humulus profilin-like protein	131	AAP15200.1	34851176	7
Humulus scandens	Japanese hop	Unassigned	Aero Plant	Humulus profilin-like protein	131	AAP15199.1	34851174	7
Juglans nigra	Black walnut	Jug n 1.0101	Food Plant	Juglans Jug r 1 Jug n 1	161	AAM54365.1	31321942	7
Juglans nigra	Black walnut	Jug n 2.0101	Food Plant	Juglans Jug r 2	481	AAM54366.1	31321944	7
Juglans nigra	Black walnut	Jug n 4.0101	Unassigned	Juglans nigra Jug n 4 legumin	510	APR62629.1	1126299828	18
Juglans regia	English walnut	Jug r 1.0101	Food Plant	Juglans Jug r 1 Jug n 1	139	AAB41308.1	1794252	7
Juglans regia	English walnut	Jug r 2.0101	Food Plant	Juglans Jug r 2	593	AAF18269.1	6580762	7
Juglans regia	English walnut	Unassigned	Food Plant	Juglans Jug r 3	119	ACI47547.1	209484145	11
Juglans regia	English walnut	Jug r 4.0101	Food Plant	Juglans Jug r 4 seed storage protein	507	AAW29810.1	56788031	7
Juniperus ashei	Mountain cedar	Jun a 2	Aero Plant	Juniperus Jun a 2	507	CAC05582.1	9955725	7
Juniperus ashei	Mountain cedar	Jun a 3.0101	Aero Plant	Juniperus Jun a 3	225	P81295.1	9087177	8
Juniperus ashei	Mountain cedar	Jun a 1.010101	Aero Plant	Juniperus Jun a/v 1	367	AAD03608.1	4138877	7
Juniperus oxycedrus	Juniper	Unassigned	Aero Plant	Juniperus Jun a/v 1	367	CAC48400.1	15139849	7
Juniperus oxycedrus	Juniper	Jun o 4	Aero Plant	Juniperus Jun o 4	165	AAC15474.2	5391446	7
Juniperus rigida	Cedar	Unassigned	Aero Plant	Juniperus Jun a 3	225	AAR21072.1	38456224	7
Juniperus rigida	Cedar	Unassigned	Aero Plant	Juniperus Jun a 3	225	AAR21071.1	38456222	7
Juniperus virginiana	Red cedar	Unassigned	Aero Plant	Juniperus Jun a 3	110	Q9LD79.2	51316532	7
Juniperus virginiana	Red cedar	Jun v 1.0102	Aero Plant	Juniperus Jun a/v 1	367	AAF80164.1	8843917	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Juniperus virginiana	Red cedar	Jun v 1.0101	Aero Plant	Juniperus Jun a/v 1	367	AAF80166.1	8843921	7
Lates calcarifer	Asian Seabass	Lat c 1.0101	Food Animal	Lates Lat c 1	109	AAV97933.1	56553743	15
Lates calcarifer	Asian Seabass	Lat c 1.0201	Food Animal	Lates Lat c 1	109	AAT45383.1	48526356	15
Lens culinaris	Lentil	Len c 3.0101	Food Plant	Lens Len c 3	118	AAX35807.1	60735410	15
Lens culinaris	Lentil	Len c 1.0101	Food Plant	Lens Len c 1	418	CAD87730.1	29539109	7
Lens culinaris	Lentil	Len c 1.0102	Food Plant	Lens Len c 1	415	CAD87731.1	29539111	7
Lepidoglyphus destructor	Storage mite	Unassigned	Aero Mite	Blomia Blo t 12	143	AAQ55550.1	33943777	7
Lepidoglyphus destructor	Storage mite	Lep d 10.0101	Aero Mite	Lepidoglyphus Lep d 10	284	CAB71342.1	6900304	15
Lepidoglyphus destructor	Storage mite	Lep d 13.0101	Aero Mite	Lepidoglyphus Lep d 13	131	CAB62213.1	6523380	15
Lepidoglyphus destructor	Storage mite	Lep d 2.0102	Aero Mite	Lepidoglyphus Lep d 2	141	CAD32313.1	21213898	7
Lepidoglyphus destructor	Storage mite	Lep d 2.0202	Aero Mite	Lepidoglyphus Lep d 2	141	CAD32314.1	21213900	7
Lepidoglyphus destructor	Storage mite	Lep d 2	Aero Mite	Lepidoglyphus Lep d 2	141	2118249B	1582223	7
Lepidoglyphus destructor	Storage mite	Lep d 2	Aero Mite	Lepidoglyphus Lep d 2	141	2118249A	1582222	7
Lepidoglyphus destructor	Storage mite	Unassigned	Aero Mite	Lepidoglyphus Lep d 2	141	AAQ73484.1	34495274	7
Lepidoglyphus destructor	Storage mite	Unassigned	Aero Mite	Lepidoglyphus Lep d 2	141	AAQ73486.1	34495278	7
Lepidoglyphus destructor	Storage mite	Unassigned	Aero Mite	Lepidoglyphus Lep d 2	140	AAQ73487.1	34495280	7
Lepidoglyphus destructor	Storage mite	Unassigned	Aero Mite	Lepidoglyphus Lep d 2	141	AAQ73488.1	34495282	7
Lepidoglyphus destructor	Storage mite	Unassigned	Aero Mite	Lepidoglyphus Lep d 2	141	AAQ73489.1	34495284	7
Lepidoglyphus destructor	Storage mite	Unassigned	Aero Mite	Lepidoglyphus Lep d 2	141	AAQ73490.1	34495286	7
Lepidoglyphus destructor	Storage mite	Unassigned	Aero Mite	Lepidoglyphus Lep d 2	141	AAQ73491.1	34495288	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Lepidoglyphus destructor	Storage mite	Unassigned	Aero Mite	Lepidoglyphus Lep d 2	141	AAQ73492.1	34495290	7
Lepidoglyphus destructor	Storage mite	Lep d 2.0101	Aero Mite	Lepidoglyphus Lep d 2	98	CAA57160.1	587450	15
Lepidoglyphus destructor	Storage mite	Lep d 2.0201	Aero Mite	Lepidoglyphus Lep d 2	141	CAA58755.1	999458	15
Lepidoglyphus destructor	Storage mite	Lep d 5.0102	Aero Mite	Lepidoglyphus Lep d 5	171	AAQ73493.1	34495292	7
Lepidoglyphus destructor	Storage mite	Lep d 5.0103	Aero Mite	Lepidoglyphus Lep d 5	169	AAQ73494.1	34495294	7
Lepidoglyphus destructor	Storage mite	Lep d 5.0101	Aero Mite	Lepidoglyphus Lep d 5	110	CAB62212.1	6523378	15
Lepidoglyphus destructor	Storage mite	Lep d 7.0101	Aero Mite	Lepidoglyphus Lep d 7	216	CAB65963.1	6706282	15
Lepidorhombus whiffiagonis	Flat fish	Lep w 1.0101	Food Animal	Lepidorhombus Lep w 1 parvalbumin	109	CAP17694.1	208608078	10
Lepisma saccharina	Silverfish	Lep s 1	Aero Insect	Lepisma Tropomyosin	284	CAC84590.2	20387027	7
Lepisma saccharina	Silverfish	Unassigned	Aero Insect	Lepisma Tropomyosin	243	CAC84593.2	20387029	7
Ligustrum vulgare	Privet	Lig v 1.0101	Aero Plant	Ligustrum Lig v 1	145	CAA54818.1	3256210	7
Ligustrum vulgare	Privet	Lig v 1.0102	Aero Plant	Ligustrum Lig v 1	145	CAA54819.1	3256212	7
Lilium longiflorum	Trumpet lily	Unassigned	Aero Plant	Lilium polygalacturonase	413	AAZ91659.1	73913442	8
Liposcelis bostrychophila	booklice	Lip b 1.0101	Unassigned	Liposcelis Lip b 1 Fragments	254	BAW03243.1	1109516247	18
Liposcelis bostrychophila	booklice	Lip b 1.0102	Unassigned	Liposcelis Lip b 1 Fragments	254	BAW03242.1	1109516245	18
Litchi chinensis	Lychee nut	Lit c 1	Food Plant	Litchi Lit c 1	131	AAL07320.1	15809696	7
Litchi chinensis	Lychee nut	Unassigned	Food Plant	Litchi Lit c 1	131	ABC02750.1	83317152	7
Litopenaeus vannamei	Whiteleg Shrimp	Lit v 4.0101	Food Animal	Litopenaeus Lit v 4 sarcoplasmic Ca+ binding	193	ACM89179.1	223403273	11
Litopenaeus vannamei	Whiteleg Shrimp	Lit v 1.0101	Food Animal	Litopenaeus Lit v 1 tropomyosin	284	ACB38288.1	170791252	10

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Litopenaeus vannamei	Whiteleg Shrimp	Lit v 2.0101	Food Animal	Litopenaeus Lit v 2	356	ABI98020.1	115492980	8
Litopenaeus vannamei	Whiteleg Shrimp	Lit v 3.0101	Food Animal	Litopenaeus Lit v 3 myosin	177	ACC76803.1	184198734	10
Lolium perenne	Perennial ryegrass	Lol p 1	Aero Plant	Lolium Lol p 1	263	P14946.2	126385	7
Lolium perenne	Perennial ryegrass	Lol p 1.0102	Aero Plant	Lolium Lol p 1	252	AAA63278.1	168314	7
Lolium perenne	Perennial ryegrass	Lol p 1.0101	Aero Plant	Lolium Lol p 1	263	AAA63279.1	168316	10
Lolium perenne	Perennial ryegrass	Lol p 1.0103	Aero Plant	Lolium Lol p 1	263	CAB63699.1	6599300	10
Lolium perenne	Perennial ryegrass	Lol p 11	Aero Plant	Lolium Lol p 11	134	Q7M1X5.1	47605808	7
Lolium perenne	Perennial ryegrass	Lol p 2.0101	Aero Plant	Lolium Lol p 2	97	P14947.1	126386	7
Lolium perenne	Perennial ryegrass	Lol p 2	Aero Plant	Lolium Lol p 2	88	CAA51775.1	939932	7
Lolium perenne	Perennial ryegrass	Lol p 3	Aero Plant	Lolium Lol p 3	97	P14948.1	126387	7
Lolium perenne	Perennial ryegrass	Lol p 4.0101	Aero Plant	Lolium Lol p 4	423	CAH92637.1	55859464	7
Lolium perenne	Perennial ryegrass	Lol p 5	Aero Plant	Lolium Lol p 5	301	AAD20386.1	4416516	7
Lolium perenne	Perennial ryegrass	Lol p 5	Aero Plant	Lolium Lol p 5	301	CAB64344.1	6634467	7
Lolium perenne	Perennial ryegrass	Lol p 5.0101	Aero Plant	Lolium Lol p 5	339	AAA33405.1	455288	10
Lolium perenne	Perennial ryegrass	Lol p 5.0102	Aero Plant	Lolium Lol p 5	307	Q40240.2	332278195	12
Lupinus albus	white lupine	Unassigned	Food Plant	Lupinus albus congluten beta	531	CAI84850.2	89994190	14
Lupinus albus	white lupine	Unassigned	Food Plant	Lupinus albus congluten beta	531	Q53HY0.2	122220821	17
Lupinus albus	white lupine	Unassigned	Food Plant	Lupinus albus congluten beta	533	Q6EBC1.1	75121065	17

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Lupinus angustifolius	blue lupin	Unassigned	Food Plant	Lupinus Lup an 1 conglutin beta	521	ABR21771.1	149208401	9
Lupinus angustifolius	blue lupin	Unassigned	Food Plant	Lupinus Lup an 1 conglutin beta	455	ABR21772.1	149208403	9
Lupinus angustifolius	blue lupin	Lup an 1.0101	Food Plant	Lupinus Lup an 1 conglutin beta	611	ACB05815.1	169950562	10
Lupinus angustifolius	blue lupin	Unassigned	Unassigned	Lupinus Lup an 1 conglutin beta	605	F5B8W5.1	980951568	17
Lupinus angustifolius	blue lupin	Unassigned	Unassigned	Lupinus Lup an 1 conglutin beta	593	F5B8W4.1	980951565	17
Lupinus angustifolius	blue lupin	Unassigned	Unassigned	Lupinus Lup an 1 conglutin beta	637	F5B8W3.1	980951561	17
Lupinus angustifolius	blue lupin	Unassigned	Unassigned	Lupinus Lup an 1 conglutin beta	590	F5B8W2.1	980951555	17
Lupinus angustifolius	blue lupin	Unassigned	Unassigned	Lupinus Lup an 1 conglutin beta	580	F5B8W1.1	980951550	17
Lupinus angustifolius	blue lupin	Unassigned	Unassigned	Lupinus Lup an 1 conglutin beta	603	F5B8W0.1	980951548	17
Lupinus angustifolius	blue lupin	Unassigned	Unassigned	Lupinus Lup an 1 conglutin beta	611	F5B8V9.1	980951518	17
Lycium barbarum	wolfberry	Unassigned	Food Plant	Lycium ltp	51	B3A0N2.1	363805423	13
Macrobrachium rosenbergii	Giant River Prawn	Mac r 1.0101	Food Animal	Macrobrachium rosenbergii shrimp tropomyosin	284	ADC55380.1	288819271	11
Macrobrachium rosenbergii	Giant River Prawn	Unassigned	Food Animal	Macrobrachium rosenbergii shrimp tropomyosin	284	AHA85706.1	558698675	15
Macruronus magellanicus	Patagonian Grenadier	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	98	P86739.1	308191450	12
Macruronus magellanicus	Patagonian Grenadier	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86741.1	308191461	12
Macruronus magellanicus	Patagonian Grenadier	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	74	P86740.1	308191474	12
Macruronus novaezelandiae	Blue hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	83	P86742.1	308191475	12
Malassezia furfur	Yeast	Mala f 2	Contact	Malassezia Mala f 2	177	BAA32435.1	3445490	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Malassezia furfur	Yeast	Mala f 3	Contact	Malassezia Mala f 3	166	BAA32436.1	3445492	7
Malassezia furfur	Yeast	Mala f 4	Contact	Malassezia Mala f 4	342	AAD25927.1	4587985	7
Malassezia sympodialis	Yeast	Mala s 1	Contact	Malassezia Mala s 1	350	CAA65341.1	1261972	7
Malassezia sympodialis	Yeast	Mala s 10.0101	Contact	Malassezia Mala s 10 heat shock protein	773	CAD20981.3	28564467	14
Malassezia sympodialis	Yeast	Mala s 11	Contact	Malassezia Mala s 11 first 38 aa signal	237	CAD68071.1	28569698	7
Malassezia sympodialis	Yeast	Mala s 12.0101	Contact	Malassezia Mala s 12	618	CAI43283.4	78038796	7
Malassezia sympodialis	Yeast	Mala s 5	Contact	Malassezia Mala s 5	172	CAA09883.1	4138171	7
Malassezia sympodialis	Yeast	Mala s 6	Contact	Malassezia Mala s 6	162	CAA09884.1	4138173	7
Malassezia sympodialis	Yeast	Mala s 7	Contact	Malassezia Mala s 7	187	CAA09885.1	4138175	7
Malassezia sympodialis	Yeast	Mala s 8	Contact	Malassezia Mala s 8	179	CAA09886.2	7271239	7
Malassezia sympodialis	Yeast	Mala s 9	Contact	Malassezia Mala s 9	342	CAA09887.4	19069920	7
Malassezia sympodialis ATCC 42132	Yeast	Unassigned	Contact	Malassezia Mala s 10 heat shock protein	773	CCU97864.1	465797105	14
Malassezia sympodialis ATCC 42132	Yeast	Unassigned	Contact	Malassezia Mala s 11 first 38 aa signal	202	CCV00099.1	465795607	14
Malassezia sympodialis ATCC 42132	Yeast	Mala s 13	Contact	Malassezia Mala s 13 Thioredoxin Rev	107	CCU98198.1	465793078	14
Malassezia sympodialis ATCC 42132	Yeast	Unassigned	Contact	Malassezia Mala s 5	172	CCU99457.1	465794772	14
Malassezia sympodialis ATCC 42132	Yeast	Unassigned	Unassigned	Malassezia Mala s 7	200	SHO79205.1	1129841119	18

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Malassezia sympodialis ATCC 42132	Yeast	Unassigned	Contact	Malassezia Mala s 9	342	CCU99206.1	465794420	14
Malus x domestica	Apple	Mal d 1.0301	Food Plant	Malus Mal d 1	159	CAA96534.1	1313966	7
Malus x domestica	Apple	Mal d 1.0401	Food Plant	Malus Mal d 1	160	CAA96535.1	1313968	7
Malus x domestica	Apple	Mal d 1.0402	Food Plant	Malus Mal d 1	160	CAA96536.1	1313970	7
Malus x domestica	Apple	Mal d 1.0403	Food Plant	Malus Mal d 1	160	CAA96537.1	1313972	7
Malus x domestica	Apple	Mal d 1.0206	Food Plant	Malus Mal d 1	159	AAD13683.1	2443824	7
Malus x domestica	Apple	Mal d 1.0103	Food Plant	Malus Mal d 1	159	AAD26546.1	4590364	7
Malus x domestica	Apple	Mal d 1.0203	Food Plant	Malus Mal d 1	159	AAD26547.1	4590366	7
Malus x domestica	Apple	Mal d 1.0204	Food Plant	Malus Mal d 1	159	AAD26548.1	4590368	7
Malus x domestica	Apple	Mal d 1.0104	Food Plant	Malus Mal d 1	159	AAD26552.1	4590376	7
Malus x domestica	Apple	Mal d 1.0105	Food Plant	Malus Mal d 1	159	AAD26553.1	4590378	7
Malus x domestica	Apple	Mal d 1.0106	Food Plant	Malus Mal d 1	159	AAD26554.1	4590380	7
Malus x domestica	Apple	Mal d 1.0107	Food Plant	Malus Mal d 1	159	AAD26555.1	4590382	7
Malus x domestica	Apple	Mal d 1.0205	Food Plant	Malus Mal d 1	159	AAD26558.1	4590388	7
Malus x domestica	Apple	Mal d 1.0208	Food Plant	Malus Mal d 1	158	CAD32318.1	21685277	7
Malus x domestica	Apple	Mal d 1.0304	Food Plant	Malus Mal d 1	159	AAO25113.1	27922941	7
Malus x domestica	Apple	Mal d 1.0108	Food Plant	Malus Mal d 1	159	AAD29671.1	4768879	11
Malus x domestica	Apple	Mal d 1.0201	Food Plant	Malus Mal d 1	159	AAB01362.1	862307	11

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Malus x domestica	Apple	Mal d 1.0102	Food Plant	Malus Mal d 1	159	CAA88833.1	886683	11
Malus x domestica	Apple	Mal d 1.0101	Food Plant	Malus Mal d 1	159	CAA58646.1	747852	15
Malus x domestica	Apple	Mal d 1.0109	Food Plant	Malus Mal d 1	159	AAK13029.1	15418742	15
Malus x domestica	Apple	Mal d 1.0207	Food Plant	Malus Mal d 1	159	AAK13030.1	15418744	15
Malus x domestica	Apple	Mal d 1.0302	Food Plant	Malus Mal d 1	159	AAK13027.1	15418738	15
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 2	26	AAB35897.1	1478293	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 2	246	AAX19848.1	60418842	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 2	246	AAX19851.1	60418848	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 2	246	Q9FSG7.1	30316292	8
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 2	158	CAT99612.1	218059718	10
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 2	158	CAT99611.1	218059715	10
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 2	193	AFM77001.1	392507603	13
Malus x domestica	Apple	Mal d 2.0101	Food Plant	Malus Mal d 2	245	AAC36740.1	3643249	15
Malus x domestica	Apple	Unassigned	Unassigned	Malus Mal d 2	159	APG29330.1	1109403341	18
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 3	115	AAT80665.1	50659891	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 3	115	AAT80664.1	50659889	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 3	115	AAT80662.1	50659885	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 3	115	AAT80659.1	50659879	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 3	115	AAT80649.1	50659859	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 3	115	AAR22488.1	38492338	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 3	115	Q9M5X7.1	14423814	9
Malus x domestica	Apple	Mal d 4.0302	Food Plant	Malus Mal d 4	131	CAD46559.1	28881453	7
Malus x domestica	Apple	Mal d 4.0102	Food Plant	Malus Mal d 4	131	CAD46561.1	28881457	7
Malus x domestica	Apple	Mal d 4.0202	Food Plant	Malus Mal d 4	131	CAD46560.1	28881455	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 4	131	AAX19854.1	60418854	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 4	131	AAX19856.1	60418858	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 4	131	AAX19858.1	60418862	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 4	131	AAX19860.1	60418866	7
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 4	131	CAK93713.1	164510842	9
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 4	131	CAK93753.1	164510858	9
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 4	131	CAK93757.1	164510860	9
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 4	77	CAT99618.1	218059730	10
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 4	115	CAT99619.1	218059733	10
Malus x domestica	Apple	Unassigned	Food Plant	Malus Mal d 4	131	CAT99617.1	218059728	10
Malus x domestica	Apple	Mal d 4.0301	Food Plant	Malus Mal d 4	131	AAD29412.1	4761584	11
Malus x domestica	Apple	Mal d 4.0201	Food Plant	Malus Mal d 4	131	AAD29413.1	4761586	11
Malus x domestica	Apple	Mal d 4.0101	Food Plant	Malus Mal d 4	131	AAD29414.1	4761588	11
Manihot esculenta	Cassava	Unassigned	Food Plant	Manihot Man e 5.0101	177	AAM55492.1	21585695	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Manihot esculenta	Cassava	Man e 5.0101	Food Plant	Manihot Man e 5.0101	177	AEE98392.1	332713934	14
Manilkara zapota	Sapodilla plum	Unassigned	Food Plant	Manilkara Thaumatin like protein 1	12	B3EWS0.1	442580988	14
Manilkara zapota	Sapodilla plum	Unassigned	Food Plant	Manilkara Thaumatin like protein 1	9	B3EWE5.3	442570282	14
Manilkara zapota	Sapodilla plum	Unassigned	Food Plant	Manilkara Thaumatin like protein 1	207	G5DC91.2	663434113	15
Marsupenaeus japonicus	Kuruma Shrimp	Unassigned	Food Animal	Marsupenaeus tropomyosin	284	BAF47263.1	125995159	8
Melicertus latisulcatus	King Prawn	Mel l 1.0101	Unassigned	Melicertus tropomyosin	284	AGF86397.1	451935062	14
Mercurialis annua	Annual mercury grass	Mer a 1	Aero Plant	Mercurialis Mer a 1	133	CAA73720.1	2959898	7
Merluccius australis australis	southern hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86745.1	308191452	12
Merluccius australis polylepis	Southern hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86749.1	308191453	12
Merluccius australis polylepis	Southern hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86750.1	308191464	12
Merluccius bilinearis	Silver hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86752.1	308191465	12
Merluccius bilinearis	Silver hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86753.1	308191478	12
Merluccius bilinearis	Silver hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	94	P86754.1	308191488	12
Merluccius capensis	Shallow-water cape hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86757.1	308191466	12
Merluccius gayi	Southern Pacific hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86761.1	308191455	12
Merluccius gayi	Southern Pacific hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	91	P86760.1	308191489	12
Merluccius merluccius	European hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P02620.1	131116	12

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Merluccius merluccius	European hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86765.1	308191469	12
Merluccius paradoxus	Deep-water cape hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86768.1	308191457	12
Merluccius paradoxus	Deep-water cape hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86769.1	308191470	12
Merluccius paradoxus	Deep-water cape hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	95	P86770.1	308191483	12
Merluccius polli	Benguela hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86771.1	308191471	12
Merluccius polli	Benguela hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	69	P86772.1	308191484	12
Merluccius productus	North Pacific hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86774.1	308191459	12
Merluccius productus	North Pacific hake	Unassigned	Food Animal	Merluccius sp. Parvalbumin Hake	108	P86775.1	308191472	12
Mesocricetus auratus	Golden hamster	Unassigned	Aero Animal	Mesocricetus auratus Mes a 1	172	AAD55792.2	13124669	16
Mesocricetus auratus	Golden hamster	Unassigned	Unassigned	Mesocricetus auratus Mes a 1	172	Q99MG7.1	81916647	17
Metapenaeus ensis	Greasyback shrimp	Met e 1	Food Animal	Metapenaeus Met e 1 Tropomyosin	274	AAA60330.1	607633	7
Mimachlamys nobilis	Noble scallop	Unassigned	Food Animal	Mimachlamys Tropomyosin	284	AAG08989.1	9954253	7
Morus alba var. atropurpurea	White Mulberry	Unassigned	Food Plant	Morus winter accumulating protein	157	AHW81906.1	610664572	15
Morus bombycis	Mulberry	Unassigned	Food Plant	Morus winter accumulating protein	157	AAV33670.1	54311115	12
Morus bombycis	Mulberry	Unassigned	Food Plant	Morus winter accumulating protein	157	AAV33672.1	54311119	12
Morus nigra	Black mulberry	Mor n 3.0101	Food Plant	Morus Mor n 3 mulberry LTP	91	P85894.1	288561913	11
Mus musculus	Mouse	Mus m 1	Aero Animal	Mus Mus m 1	180	P02762.2	20178291	7
Mus musculus	Mouse	Mus m 1.0101	Aero Animal	Mus Mus m 1	180	CAA26953.1	295910	15
Mus musculus	Mouse	Unassigned	Unassigned	Mus Mus m 1	181	A2BIM8.1	980952242	17
Mus musculus domesticus	Mouse	Mus m 1.0102	Aero Animal	Mus Mus m 1	180	AAA39768.1	199881	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Musa acuminata	Banana	Mus xp 1	Food Plant	Musa acuminata Mus a 1 profilin banana	131	AAK54834.1	14161635	7
Musa acuminata	Banana	Unassigned	Food Plant	Musa acuminata Mus a 5 Endo-Beta-1,3-Glucanase	312	2CYG_A	83754908	7
Musa acuminata	Banana	Mus a 4.0101	Food Plant	Musa Mus a 4	200	1Z3Q_A	88191901	7
Musa acuminata	Banana	Mus a 2.0101	Food Plant	Musa Mus s 2	318	CAC81811.1	17932710	15
Musa acuminata AAA Group	Banana	Unassigned	Food Plant	Musa acuminata Mus a 5 Endo-Beta-1,3-Glucanase	340	AAB82772.2	6073860	14
Myrmecia banksi	Giant Bull Ant	Myr p 3.0101	Venom or Salivary	Myrmecia Myr p 3	84	BAD36780.1	51241753	15
Myrmecia pilosula	Jumper ant	Unassigned	Venom or Salivary	Myrmecia Myr p 1	112	AAB50883.1	1911819	7
Myrmecia pilosula	Jumper ant	Myr p 1.0101	Venom or Salivary	Myrmecia Myr p 1	112	CAA49760.1	312284	15
Myrmecia pilosula	Jumper ant	Myr p 2	Venom or Salivary	Myrmecia Myr p 2	75	2206305A	1587177	7
Myrmecia pilosula	Jumper ant	Myr p 2.0101	Venom or Salivary	Myrmecia Myr p 2	75	AAB36316.1	1438761	10
Neptunea polycostata	Wrinkled Neptune	Unassigned	Food Animal	Neptunea tropomyosin	284	BAH10150.1	219806590	10
Nicotiana tabacum	Tobacco	Unassigned	Aero Plant	Nicotiana villin	520	CAE17317.1	57283139	7
Nicotiana tabacum	Tobacco	Unassigned	Aero Plant	Nicotiana villin	559	CAE17316.1	57283137	7
Octopus vulgaris	Octopus	Unassigned	Food Animal	Octopus tropomyosin	284	BAE54433.1	83715936	7
Olea europaea	Olive tree	Ole e 1	Aero Plant	Olea Ole e 1	145	P19963.2	14424429	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	137	I53806	1362128	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	136	E53806	1362129	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	136	F53806	1362130	7
Olea europaea	Olive tree	Ole e 1.0104	Aero Plant	Olea Ole e 1	145	C53806	1362131	7
Olea europaea	Olive tree	Ole e 1	Aero Plant	Olea Ole e 1	137	A38968	1362132	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	136	G53806	1362133	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Olea europaea	Olive tree	Ole e 1.0103	Aero Plant	Olea Ole e 1	145	B53806	1362136	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	136	H53806	1362137	7
Olea europaea	Olive tree	Ole e 1.0105	Aero Plant	Olea Ole e 1	146	CAA73038.1	2465127	7
Olea europaea	Olive tree	Ole e 1.0106	Aero Plant	Olea Ole e 1	146	CAA73037.1	2465129	7
Olea europaea	Olive tree	Ole e 1.0107	Aero Plant	Olea Ole e 1	146	CAA73036.1	2465131	7
Olea europaea	Olive tree	Ole e 1.0101	Aero Plant	Olea Ole e 1	130	AAB32652.2	13195753	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	134	AAO22133.1	37724597	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	135	AAO22132.1	37724593	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	132	AAN18044.1	37548753	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	131	AAQ10281.1	33329758	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	132	AAQ10280.1	33329756	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	132	AAQ10279.1	33329754	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	131	AAQ10278.1	33329752	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	131	AAQ10277.1	33329750	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	129	AAQ10276.1	33329748	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	131	AAQ10274.1	33329744	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	132	AAQ10271.1	33329738	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	132	AAQ10268.1	33329732	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	132	AAQ08190.1	33325115	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	140	ABP58632.1	145313982	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	140	ABP58633.1	145313984	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	140	ABP58635.1	145313988	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	140	ABP58636.1	145313990	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 1	140	ABP58637.1	145313992	9
Olea europaea	Olive tree	Ole e 10	Aero Plant	Olea Ole e 10	123	AAL92578.1	29465664	7
Olea europaea	Olive tree	Ole e 11.0102	Aero Plant	Olea Ole e 11.0101 and 0102	364	AAY88919.1	68270856	11
Olea europaea	Olive tree	Ole e 11.0101	Aero Plant	Olea Ole e 11.0101 and 0102	364	ACZ57582.1	269996495	11
Olea europaea	Olive tree	Unassigned	Aero Plant	olea Ole e 12	308	E1U332.1	449061782	14
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 13	226	E3SU11.1	449061783	14

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Olea europaea	Olive tree	Ole e 2	Aero Plant	Olea Ole e 2	134	O24170.1	3914427	7
Olea europaea	Olive tree	Ole e 2	Aero Plant	Olea Ole e 2	134	O24171.1	3914428	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 2	131	A4GFC0.1	576017874	15
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 2	131	A4GFC3.1	576017774	15
Olea europaea	Olive tree	Ole e 2.0101	Aero Plant	Olea Ole e 2	134	CAA73035.1	2465133	15
Olea europaea	Olive tree	Ole e 3.0101	Aero Plant	Olea Ole e 3	84	AAD05375.1	3337403	7
Olea europaea	Olive tree	Ole e 3	Aero Plant	Olea Ole e 3	52	AAO33897.1	37725377	7
Olea europaea	Olive tree	Ole e 5.0101	Aero Plant	Olea Ole e 5	30	P80740.2	122064581	8
Olea europaea	Olive tree	Ole e 5	Aero Plant	Olea Ole e 5	152	CAD21706.2	39840779	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABP58627.1	145313972	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX26131.1	160347106	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	144	ABX26132.1	160347108	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX26134.1	160347112	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX26138.1	160347120	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX26139.1	160347122	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX26140.1	160347124	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX26141.1	160347126	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX26143.1	160347130	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX26145.1	160347134	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX26147.1	160347138	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX54842.1	160962543	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX54844.1	160962547	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX54849.1	160962557	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX54855.1	160962569	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX54859.1	160962577	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX54862.1	160962583	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	144	ABX54864.1	160962587	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX54866.1	160962591	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX54869.1	160962597	9
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX54876.1	160962611	9

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 5	152	ABX54877.1	160962613	9
Olea europaea	Olive tree	Ole e 6.0101	Aero Plant	Olea Ole e 6	50	AAB66909.1	2276458	11
Olea europaea	Olive tree	Ole e 7	Aero Plant	Olea Ole e 7	21	P81430.2	22002032	7
Olea europaea	Olive tree	Ole e 8	Aero Plant	Olea Ole e 8	171	AAF31152.1	6901654	7
Olea europaea	Olive tree	Ole e 8.0101	Aero Plant	Olea Ole e 8	171	AAF31151.1	6901652	11
Olea europaea	Olive tree	Ole e 9	Aero Plant	Olea Ole e 9	460	AAK58515.1	14279169	7
Olea europaea	Olive tree	Unassigned	Aero Plant	Olea Ole e 9	101	2JON_A	166235350	9
Ommastrephes bartramii	red squid	Unassigned	Food Animal	Ommastrephes tropomyosin	284	BAE54432.1	83715934	7
Onchocerca volvulus	Parasitic nematode	Unassigned	Worm (parasite)	Onchocerca tropomyosin	284	Q25632.1	42559586	12
Oncorhynchus keta	chum salmon	Onc k 5.0101	Food Animal	Oncorhynchus Onc k 5	193	BAJ07603.1	296040357	15
Oncorhynchus mykiss	rainbow trout	Onc m 1.0101	Food Animal	Oncorhynchus Rainbow trout parv Onc m 1	108	P86431.1	288559139	11
Oncorhynchus mykiss	rainbow trout	Onc m 1.0201	Food Animal	Oncorhynchus Rainbow trout parv Onc m 1	107	P86432.1	288559140	11
Oratosquilla oratoria	mantis shrimp	Unassigned	Food Animal	Oratosquilla tropomyosin	284	BAF95206.1	162286975	9
Oreochromis mossambicus	Mozambique tilapia	Ore m 4.0101	Food Animal	Oreochromis Ore m 4 tropomyosin	284	AFV53352.1	410060781	14
Oryctolagus cuniculus	European rabbit	Ory c 3.A.0101	Aero Animal	Oryctolagus Ory c 3	93	AAG42806.1	11993600	15
Oryctolagus cuniculus	pistachio	Ory c 3.B.0101	Aero Animal	Oryctolagus Ory c 3	90	AAG42802.1	11993592	15
Oryza sativa	Rice	Unassigned	Food Plant	Oryza Glyoxalase I	291	Q948T6.2	84029333	7
Oryza sativa	Rice	Ory s 1.0101	Aero Plant	Oryza Ory s 1	263	AAA86533.1	1173557	8
Oryza sativa	Rice	Unassigned	Aero Plant	Oryza Ory s 1	267	AAF72991.1	8118439	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Glyoxalase I	291	BAB71741.1	16580747	7
Oryza sativa (japonica cultivar-group)	Rice	Ory s 1	Aero Plant	Oryza Ory s 1	267	Q40638.2	109913547	8

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Aero Plant	Oryza putative polcalcin Phl p 7	82	BAD13150.1	45736119	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	157	BAC20657.1	23616954	8
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	165	BAA01998.1	218193	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	157	BAA01996.1	218197	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	111	BAA07772.1	1304216	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	109	BAA07773.1	1304217	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	113	BAA07774.1	1304218	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	166	BAA07710.1	1398913	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	160	BAA07711.1	1398915	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	157	BAA07712.1	1398916	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	160	BAA07713.1	1398918	7
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	157	AAB99797.1	2827316	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	166	Q01882.2	114152865	8
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	163	Q01883.2	114152864	8
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	160	BAC19997.1	23495787	8
Oryza sativa (japonica cultivar-group)	Rice	Unassigned	Food Plant	Oryza Trypsin alpha-amylase inhibitor	160	BAC20650.1	23616947	7
Ostrya carpinifolia	European hop hornbeam	Ost c 1.0101	Aero Plant	Ostrya Ost c 1pollen allergen	160	ADK39021.1	300872535	12
Pachycondyla chinensis	Asian needle ant	Unassigned	Venom or Salivary	Pachycondyla Pac c 3 allergen	199	ACA96507.1	169822894	10
Pandalus borealis	caribean shrimp	Pan b 1.0101	Food Animal	Pandalus Pan b 1	284	CBY17558.1	312831088	12
Panulirus stimpsoni	Lobster	Pan s 1.0101	Food Animal	Panulirus Pan s 1	274	AAC38996.1	3080761	11
Paralithodes camtschaticus	Kamchatka crab	Unassigned	Food Animal	Paralithodes tropomyosin	284	BAF47265.1	125995163	8
Paralithodes camtschaticus	Kamchatka crab	Unassigned	Food Animal	Paralithodes tropomyosin	284	BAF47266.1	125995165	8
Parietaria judaica	Weed	Par j 1	Aero Plant	Parietaria Par j 1	143	2008179A	741844	7
Parietaria judaica	Weed	Par j 1.0102	Aero Plant	Parietaria Par j 1	176	CAA65123.1	1532058	7
Parietaria judaica	Weed	Par j 1.0101	Aero Plant	Parietaria Par j 1	133	CAA54587.1	992612	15
Parietaria judaica	Weed	Par j 1.0103	Aero Plant	Parietaria Par j 1	139	CAI94601.1	95007033	15
Parietaria judaica	Weed	Par j 1.0201	Aero Plant	Parietaria Par j 1	138	CAA59370.1	706811	15
Parietaria judaica	Weed	Par j 2.0102	Aero Plant	Parietaria Par j 2	133	CAA65122.1	1532056	7
Parietaria judaica	Weed	Par j 2.0101	Aero Plant	Parietaria Par j 2	133	P55958.1	2497750	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Parietaria judaica	Weed	Par j 3.0102	Aero Plant	Parietaria Par j 3 profilin	131	Q9T0M8.1	14423869	7
Parietaria judaica	Weed	Par j 3.0101	Aero Plant	Parietaria Par j 3 profilin	132	Q9XG85.1	14423876	7
Parietaria judaica	Weed	Par j 3.0201	Aero Plant	Parietaria Par j 3 profilin	131	CCP19647.1	444175753	14
Parietaria judaica	Weed	Par j 4.0101	Aero Plant	Parietaria Par j 4	84	CAP05019.1	201071363	15
Parietaria officinalis	Weed	Par o 1	Aero Plant	Parietaria Par o 1	12	Q7M1E8	75139847	7
Parietaria officinalis	Weed	Par o 1	Aero Plant	Parietaria Par o 1	17	AAB36008.1	1311509	7
Parietaria officinalis	Weed	Par o 1	Aero Plant	Parietaria Par o 1	15	AAB36009.1	1311510	7
Parietaria officinalis	Weed	Par o 1	Aero Plant	Parietaria Par o 1	15	AAB36010.1	1311511	7
Parietaria officinalis	Weed	Par o 1	Aero Plant	Parietaria Par o 1	15	AAB36011.1	1311512	7
Parietaria officinalis	Weed	Par o 1	Aero Plant	Parietaria Par o 1	30	AAB36012.1	1311513	7
Parietaria officinalis	Weed	Par o 1	Aero Plant	Parietaria Par o 1	24	AAB46820.1	1836011	7
Parietaria officinalis	Weed	Unassigned	Aero Plant	Parietaria Par o 1	25	AAB46819.1	1836010	7
Parthenium hysterophorus	Pollen defensin	Unassigned	Unassigned	Parthenium hysterophorus Par h 1	156	AKF12278.1	817033923	17
Paspalum notatum	Bahia grass	Unassigned	Aero Plant	Paspalum group 13 pollen allergen	169	CBM42667.1	338930686	12
Paspalum notatum	Bahia grass	Unassigned	Aero Plant	Paspalum group 13 pollen allergen	169	CBM42666.1	338930684	12
Paspalum notatum	Bahia grass	Unassigned	Aero Plant	Paspalum group 13 pollen allergen	169	CBM42665.1	338930682	12
Paspalum notatum	Bahia grass	Unassigned	Aero Plant	Paspalum group 13 pollen allergen	169	CBM42664.1	338930680	12
Paspalum notatum	Bahia grass	Unassigned	Aero Plant	Paspalum group 13 pollen allergen	393	CBM42663.1	338930678	12
Paspalum notatum	Bahia grass	Unassigned	Aero Plant	Paspalum group 13 pollen allergen	393	CBM42662.1	338930676	12

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Paspalum notatum	Bahia grass	Unassigned	Aero Plant	Paspalum group 13 pollen allergen	391	CBM42661.1	338930674	12
Paspalum notatum	Bahia grass	Unassigned	Aero Plant	Paspalum group 13 pollen allergen	395	CBM42660.1	338930672	12
Paspalum notatum	Bahia grass	Pas n 1.0101	Aero Plant	Paspalum Pas n 1 beta expansin	265	ACA23876.1	168419914	10
Penaeus monodon	Black tiger shrimp	Pen m 1.0101	Food Animal	Penaeus Pen m 1 tropomyosin	284	AAX37288.1	60892782	15
Penaeus monodon	Black tiger shrimp	Pen m 2	Food Animal	Penaeus Pen m 2	356	AAO15713.1	27463265	7
Penaeus monodon	Black tiger shrimp	Unassigned	Food Animal	Penaeus Pen m 2	356	C7E3T4.1	308154236	12
Penaeus monodon	Black tiger shrimp	Pen m 3.0101	Food Animal	Penaeus Pen m 3 myosin light chain	177	ADV17342.1	317383196	12
Penaeus monodon	Black tiger shrimp	Pen m 4.0101	Food Animal	Penaeus Pen m 4 sarcoplasmic calcium binding	193	ADV17343.1	317383198	12
Penicillium brevicompactum	Fungus	Pen b 26.0101	Aero Fungi	Penicillium Pen b 26	107	AAX11194.1	59894749	7
Penicillium chrysogenum	Fungus	Pen ch 18.0101	Aero Fungi	Penicillium Pen 18	494	AAF71379.1	7963902	7
Penicillium chrysogenum	Fungus	Pen ch 18	Aero Fungi	Penicillium Pen 18	494	AAG44693.2	14215732	7
Penicillium chrysogenum	Fungus	Pen ch 13.0101	Aero Fungi	Penicillium Pen ch 13	397	AAF23726.1	6684758	7
Penicillium chrysogenum	Fungus	Pen ch 13	Aero Fungi	Penicillium Pen ch 13	398	AAM33821.1	21069093	7
Penicillium chrysogenum	Fungus	Pen ch 20	Aero Fungi	Penicillium Pen ch 20	117	AAB34785.1	999009	7
Penicillium chrysogenum	Fungus	Pen ch 35.0101	Aero Fungi	Penicillium Pen ch 35	324	ADK27483.1	300679427	15
Penicillium citrinum	Fungus	Unassigned	Aero Fungi	Penicillium Pen 18	457	AAD25995.1	4588118	7
Penicillium citrinum	Fungus	Unassigned	Aero Fungi	Penicillium Pen 18	358	AAG44480.1	12005501	7
Penicillium citrinum	Fungus	Pen c 19	Aero Fungi	Penicillium Pen c 19	503	Q92260.1	14423733	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Penicillium citrinum	Fungus	Pen c 22	Aero Fungi	Penicillium Pen c 22	438	AAK51201.1	13991101	7
Penicillium citrinum	Fungus	Pen c 24	Aero Fungi	Penicillium Pen c 24	228	AAR17475.1	38326693	7
Penicillium citrinum	Fungus	Pen c 3	Aero Fungi	Penicillium Pen c 3	167	AAD42074.1	5326864	7
Penicillium citrinum	Fungus	Pen c 30.0101	Aero Fungi	Penicillium Pen c 30	733	ABB89950.1	82754305	7
Penicillium citrinum	Fungus	Pen c 32.0101	Aero Fungi	Penicillium Pen c 32	290	ABM60783.1	121584258	8
Penicillium citrinum	Fungus	Unassigned	Aero Fungi	Penicillium Pen ch 13	397	AAD25926.1	4587983	7
Penicillium crustosum	Fungus	Pen cr 26.0101	Aero Fungi	Penicillium crustosum Pen cr 26 60s P1	107	AEX34122.1	371537645	13
Penicillium oxalicum	Fungus	Pen o 18.0101	Aero Fungi	Penicillium Pen 18	503	AAG44478.1	12005497	7
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta americana Per a 11	494	AKH04310.1	821092692	16
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta americana Per a 12	407	AKH04311.1	821092694	16
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta GST Per a 5	216	AAX33729.1	60678789	7
Periplaneta americana	American cockroach	Per a 5.0102	Aero Insect	Periplaneta GST Per a 5	216	AEV23867.1	359326557	15
Periplaneta americana	American cockroach	Per a 7.0102	Aero Insect	Periplaneta Per 7	284	AAD19606.1	4378573	7
Periplaneta americana	American cockroach	Per a 7.0101	Aero Insect	Periplaneta Per 7	284	CAB38086.1	4468639	7
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta Per 7	284	ACS14052.1	239740599	11
Periplaneta americana	American cockroach	Per a 1.0201	Aero Insect	Periplaneta Per a 1	446	AAC34736.1	2231297	7
Periplaneta americana	American cockroach	Per a 1.0104	Aero Insect	Periplaneta Per a 1	274	AAC34737.1	2253610	7
Periplaneta americana	American cockroach	Per a 1.0103	Aero Insect	Periplaneta Per a 1	395	AAB82404.1	2580504	7
Periplaneta americana	American cockroach	Per a 1.0102	Aero Insect	Periplaneta Per a 1	228	AAC34312.1	2897849	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Periplaneta americana	American cockroach	Per a 1.0101	Aero Insect	Periplaneta Per a 1	231	AAD13533.1	4240399	7
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta Per a 1	124	AAP13554.1	30144660	7
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta Per a 1	395	ADB92492.1	284518361	11
Periplaneta americana	American cockroach	Per a 10.0101	Aero Insect	Periplaneta Per a 10 ser protease	256	AAX33734.1	60678799	7
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta Per a 2	351	AAX33727.1	60678785	7
Periplaneta americana	American cockroach	Per a 2.0101	Aero Insect	Periplaneta Per a 2	351	ADR82198.1	313870534	12
Periplaneta americana	American cockroach	Per a 3.0201	Aero Insect	Periplaneta Per a 3	631	AAB09632.1	1531589	7
Periplaneta americana	American cockroach	Per a 3.0202	Aero Insect	Periplaneta Per a 3	470	AAB62731.1	1580794	7
Periplaneta americana	American cockroach	Per a 3.0203	Aero Insect	Periplaneta Per a 3	393	AAB63595.1	1580797	7
Periplaneta americana	American cockroach	Per a 3.0101	Aero Insect	Periplaneta Per a 3	685	Q25641.1	2833325	9
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta Per a 3	688	ADB92493.1	284518363	11
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta Per a 3	685	ADD17628.1	289721058	11
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta Per a 4	183	AAX33728.1	60678787	7
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta Per a 4	163	3EBW_A	215794707	10
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta Per a 4	167	ACJ37391.1	212675312	10
Periplaneta americana	American cockroach	Per a 5.0102	Aero Insect	Periplaneta Per a 6	151	AAX33730.1	60678791	8
Periplaneta americana	American cockroach	Unassigned	Aero Insect	Periplaneta Per a 9	356	AAT77152.1	50428904	8
Periplaneta americana	American cockroach	Per a 9.0101	Aero Insect	Periplaneta Per a 9	356	ACA00204.1	167782135	9
Periplaneta fuliginosa	Smokybrown cockroach	Unassigned	Aero Insect	Periplaneta Per 7	284	AAL86701.1	19310971	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Perna viridis	Asian green mussell	Unassigned	Food Animal	Perna Tropomyosin	284	AAG08988.1	9954251	7
Persea americana	Avocado	Pers a 1	Food Plant	Persea Pers a 1	326	CAB01591.1	3201547	7
Phalaris aquatica	Canary grass	Unassigned	Aero Plant	Phalaris Pha a 1	20	AAB27445.1	409328	7
Phalaris aquatica	Canary grass	Pha a 1	Aero Plant	Phalaris Pha a 1	269	Q41260.1	2498576	7
Phalaris aquatica	Canary grass	Pha a 5.0101	Aero Plant	Phalaris Pha a 5	320	P56164.1	2498577	7
Phalaris aquatica	Canary grass	Unassigned	Aero Plant	Phalaris Pha a 5	305	P56165.1	2498578	7
Phalaris aquatica	Canary grass	Unassigned	Aero Plant	Phalaris Pha a 5	294	P56166.1	2498579	7
Phalaris aquatica	Canary grass	Unassigned	Aero Plant	Phalaris Pha a 5	175	P56167.1	2498580	7
Phaseolus vulgaris	Kidney bean	Pha v 3.0101	Food Plant	Phaseolus Pha v 3	115	ADC80502.1	289064177	11
Phaseolus vulgaris	Kidney bean	Pha v 3.0201	Food Plant	Phaseolus Pha v 3	118	ADC80503.1	289064179	11
Phleum pratense	Common timothy	Phl p 1.0102	Aero Plant	Phleum Phl p 1	263	CAA55390.1	473360	7
Phleum pratense	Common timothy	Phl p 1.0101	Aero Plant	Phleum Phl p 1	263	CAA81613.1	3901094	7
Phleum pratense	Common timothy	Phl p 1	Aero Plant	Phleum Phl p 1	241	1N10_A	28373838	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 1	240	CAG24374.1	45823012	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 1	262	2118271A	1582250	10
Phleum pratense	Common timothy	Phl p 11	Aero Plant	Phleum Phl p 11	143	AAN32987.1	23452313	7
Phleum pratense	Common timothy	Phl p 12.0103	Aero Plant	Phleum Phl p 12	131	CAA70609.1	2415700	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 12	131	ABG81289.1	110644906	8
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 12	131	ABG81290.1	110644908	8

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 12	131	ABG81291.1	110644910	8
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 12	131	ABG81292.1	110644912	8
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 12	131	ABG81293.1	110644914	8
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 12	131	ABG81294.1	110644916	8
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 12	131	ABG81295.1	110644918	8
Phleum pratense	Common timothy	Phl p 12.0102	Aero Plant	Phleum Phl p 12	131	CAA70608.1	2415698	10
Phleum pratense	Common timothy	Phl p 12.0101	Aero Plant	Phleum Phl p 12	131	CAA54686.1	453976	15
Phleum pratense	Common timothy	Phl p 13	Aero Plant	Phleum Phl p 13	394	CAB42886.1	4826572	7
Phleum pratense	Common timothy	Phl p 2	Aero Plant	Phleum Phl p 2	122	CAA53529.1	415896	7
Phleum pratense	Common timothy	Phl p 4.0101	Aero Plant	Phleum Phl p 4	508	CAD54670.2	54144332	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 4	500	CAF32567.2	45108973	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 4	500	CAF32566.2	45108967	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 4	500	CAQ55938.1	189014266	10
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 4	500	CAQ55939.1	189014268	10
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 4	500	CAQ55940.1	189014270	10
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 4	500	CAQ55941.1	189014272	10
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 4	500	3TSH_A	405944794	14
Phleum pratense	Common timothy	Phl p 4.0201	Aero Plant	Phleum Phl p 4	508	CAD54671.2	54144334	15
Phleum pratense	Common timothy	Phl p 5.0101	Aero Plant	Phleum Phl p 5	312	CAA52753.1	398830	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Phleum pratense	Common timothy	Phl p 5	Aero Plant	Phleum Phl p 5	257	S32101	422005	7
Phleum pratense	Common timothy	Phl p 5	Aero Plant	Phleum Phl p 5	280	S38584	481397	7
Phleum pratense	Common timothy	Phl p 5	Aero Plant	Phleum Phl p 5	24	Q7M1L8	75139900	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	285	2023228A	1092249	7
Phleum pratense	Common timothy	Phl p 5.0202	Aero Plant	Phleum Phl p 5	281	CAB05371.1	1684718	7
Phleum pratense	Common timothy	Phl p 5.0104	Aero Plant	Phleum Phl p 5	276	CAB05372.1	1684720	7
Phleum pratense	Common timothy	Phl p 5.0102	Aero Plant	Phleum Phl p 5	286	CAA50281.1	2398757	7
Phleum pratense	Common timothy	Phl p 5.0105	Aero Plant	Phleum Phl p 5	276	AAC16525.1	3135497	7
Phleum pratense	Common timothy	Phl p 5.0106	Aero Plant	Phleum Phl p 5	276	AAC16526.1	3135499	7
Phleum pratense	Common timothy	Phl p 5.0107	Aero Plant	Phleum Phl p 5	276	AAC16527.1	3135501	7
Phleum pratense	Common timothy	Phl p 5.0108	Aero Plant	Phleum Phl p 5	276	AAC16528.1	3135503	7
Phleum pratense	Common timothy	Phl p 5.0103	Aero Plant	Phleum Phl p 5	312	AAC25994.1	3309039	7
Phleum pratense	Common timothy	Phl p 5.0203	Aero Plant	Phleum Phl p 5	295	AAC25995.1	3309041	7
Phleum pratense	Common timothy	Phl p 5.0206	Aero Plant	Phleum Phl p 5	290	AAC25997.1	3309045	7
Phleum pratense	Common timothy	Phl p 5.0207	Aero Plant	Phleum Phl p 5	287	AAC25998.1	3309047	7
Phleum pratense	Common timothy	Phl p 5	Aero Plant	Phleum Phl p 5	275	AAK25823.1	13430402	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38384.1	21725606	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38385.1	21725608	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38386.1	21725610	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38387.1	21725612	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38388.1	21725614	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38389.1	21725616	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38390.1	21725618	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38391.1	21725620	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38392.1	21725622	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38393.1	21725624	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38394.1	21725626	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38395.1	21725628	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38396.1	21725630	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	287	CAD38397.1	21725632	7
Phleum pratense	Common timothy	Phl p 5	Aero Plant	Phleum Phl p 5	102	1L3P_A	28948464	7
Phleum pratense	Common timothy	Phl p 5.0109	Aero Plant	Phleum Phl p 5	284	CAD87529.1	29500897	7
Phleum pratense	Common timothy	Phl p 5.0201	Aero Plant	Phleum Phl p 5	284	CAA81609.1	2398759	10
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 5	309	CCD28287.1	345108717	13
Phleum pratense	Common timothy	Phl p 6.0102	Aero Plant	Phleum Phl p 6	138	CAA76556.1	3004465	7
Phleum pratense	Common timothy	Phl p 6.0101	Aero Plant	Phleum Phl p 6	138	CAA76557.1	3004467	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 6	106	CAA76558.1	3004469	7
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum Phl p 6	111	1NLX_N	28374072	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Phleum pratense	Common timothy	Phl p 7.0101	Aero Plant	Phleum Polcalin (Phl p 7)	78	CAA76887.1	3367732	10
Phleum pratense	Common timothy	Unassigned	Aero Plant	Phleum pollen allergen group 3	100	3FT1_A	283806867	11
Phodopus sungorus	Siberian hamster	Phod s 1.0101	Aero Plant	Phodopus sungorus lipocalin	151	AGT28425.1	530376029	16
Phoenix dactylifera	Date palm	Pho d 2	Aero Plant	Phoenix Pho d 2	131	CAD10390.1	21322677	7
Pinus koraiensis	Vicilin	Pin k 2.0101	Plant food	Pinus koraiensis vicilin Pin k 2.0101	463	AHC94918.1	567773309	15
Pinus pinea	Pine	Unassigned	Plant Food	Pinus pinea albumin AAI	110	CEJ95862.1	749495809	16
Pinus pinea	2S albumin	Pin p 1.0101	Plant food	Pinus pinea Pin p 1 2S	164	CTQ87571.1	916237486	16
Pistacia vera	Pistachio	Unassigned	Food Plant	Pistacia 11S globulin	472	ABU42022.1	156001070	9
Pistacia vera	Pistachio	Pis v 2.0101	Food Plant	Pistacia 11S globulin	496	ABG73109.1	110349083	10
Pistacia vera	Pistachio	Pis v 2.0201	Food Plant	Pistacia 11S globulin	472	ABG73110.1	110349085	10
Pistacia vera	Pistachio	Pis v 1.0101	Food Plant	Pistacia Pis v 1 2S albumin	149	ABG73108.1	110349081	10
Pistacia vera	Pistachio	Pis v 3.0101	Food Plant	Pistacia Pis v 3 vicilin	519	ABO36677.1	133711974	10
Pistacia vera	Pistachio	Pis v 4.0101	Food Plant	Pistacia Pis v 4	230	ABR29644.1	149786150	9
Pisum sativum	Pea	Pis s 1.0102	Food Plant	Pisum Pis s 1	415	CAF25233.1	42414629	7
Pisum sativum	Pea	Pis s 1.0101	Food Plant	Pisum Pis s 1	415	CAF25232.1	42414627	7
Pisum sativum	Pea	Pis s 2.0101	Food Plant	Pisum Pis s 2	613	CAB82855.1	7339551	15
Pisum sativum	Pea	Pis s 3.0101	Unassigned	Pisum sativum Pis s 3	120	AJG44053.1	752855036	17
Pisum sativum	Pea	Unassigned	Unassigned	Pisum sativum Pis s 3	120	A0A158V755.1	1064302992	18
Pisum sativum	Pea	Unassigned	Unassigned	Pisum sativum Pis s 3	119	A0A158V976.1	1064302965	18
Pisum sativum	Pea	Unassigned	Unassigned	Pisum sativum Pis s 3	95	2N81_A	1026943499	18
Plantago lanceolata	Narrow-leaved plantain	Pla l 1.0101	Aero Plant	Plantago Pla l 1	131	CAC41633.1	14422359	7
Plantago lanceolata	Narrow-leaved plantain	Pla l 1.0102	Aero Plant	Plantago Pla l 1	131	CAC41634.1	14422361	7
Plantago lanceolata	Narrow-leaved plantain	Pla l 1.0103	Aero Plant	Plantago Pla l 1	131	CAC41635.1	14422363	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Plantago lanceolata	Narrow-leaved plantain	Unassigned	Aero Plant	Plantago Pla l 1	65	CAD80019.1	29163773	7
Platanus orientalis	oriental plane	Pla or 1.0101	Aero Plant	Platanus Pla or 1	170	ABY21305.1	162949336	9
Platanus orientalis	oriental plane	Pla or 2.0101	Aero Plant	Platanus Pla or 2	378	ABY21306.1	162949338	9
Platanus x acerifolia	London plane tree	Pla a 3.0101	Aero Plant	Platanus acerifolia Pla a 3	93	ALF39466.1	930156468	16
Platanus x acerifolia	London plane tree	Pla a 3.0201	Aero Plant	Platanus acerifolia Pla a 3	118	ALF00099.1	928541035	17
Platanus x acerifolia	London plane tree	Pla a 1	Aero Plant	Platanus Pla a 1	179	CAD20556.1	26190140	7
Platanus x acerifolia	London plane tree	Pla a 2	Aero Plant	Platanus Pla a 2	377	CAE52833.1	49523394	7
Plodia interpunctella	Indian meal moth	Plo i 1.0101	Aero Insect	Plodia Plo i 1 Arginine kinase	355	CAC85911.1	15886861	7
Plodia interpunctella	Indian meal moth	Plo i 2.0101	Aero Insect	Plodia Plo i 2 thioredoxin	106	CBW45298.1	308193268	12
Poa pratensis	Kentucky bluegrass	Poa p 1	Aero Plant	Poa Poa p 1	20	A60372	280414	7
Poa pratensis	Kentucky bluegrass	Poa p 1	Aero Plant	Poa Poa p 1	26	F37396	320620	7
Poa pratensis	Kentucky bluegrass	Poa p 1.0101	Aero Plant	Poa Poa p 1	263	CAA10520.1	4090265	7
Poa pratensis	Kentucky bluegrass	Poa p 5	Aero Plant	Poa Poa p 5	303	AAG42254.1	11991227	7
Poa pratensis	Kentucky bluegrass	Unassigned	Aero Plant	Poa Poa p 9	373	P22284.1	113560	7
Poa pratensis	Kentucky bluegrass	Unassigned	Aero Plant	Poa Poa p 9	307	P22286.1	113562	7
Poa pratensis	Kentucky bluegrass	Unassigned	Aero Plant	Poa Poa p 9	131	A60373	539056	7
Poa pratensis	Kentucky bluegrass	Unassigned	Aero Plant	Poa Poa p 9	333	P22285.1	113561	7
Polistes annularis	Paper wasp	Pol a 5.0101	Venom or Salivary	Polistes Pol 5	209	AAA29793.1	160780	7
Polistes annularis	Paper wasp	Pol a 1.0101	Venom or Salivary	Polistes Pol a 1 Pol d 1	301	AAD52615.1	5815249	11

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Polistes annularis	Paper wasp	Pol a 2.0101	Venom or Salivary	Polistes Pol a 2	367	AAD52616.1	5815251	11
Polistes dominulus	Paper wasp	Pol d 5	Venom or Salivary	Polistes Pol 5	227	AAT95010.1	51093377	7
Polistes dominulus	Paper wasp	Pol d 1.0104	Venom or Salivary	Polistes Pol a 1 Pol d 1	316	AAS67044.1	45510893	7
Polistes dominulus	Paper wasp	Pol d 1.0103	Venom or Salivary	Polistes Pol a 1 Pol d 1	316	AAS67043.1	45510891	7
Polistes dominulus	Paper wasp	Pol d 1.0102	Venom or Salivary	Polistes Pol a 1 Pol d 1	316	AAS67042.1	45510889	7
Polistes dominulus	Paper wasp	Pol d 1.0101	Venom or Salivary	Polistes Pol a 1 Pol d 1	337	AAS67041.1	45510887	7
Polistes dominulus	Paper wasp	Pol d 4.0101	Venom or Salivary	Polistes Venom serine protease	277	AAP37412.1	30909091	7
Polistes exclamans	Paper wasp	Pol e 5.0101	Venom or Salivary	Polistes Pol 5	226	AAT95009.1	51093375	7
Polistes fuscatus	Paper wasp	Pol f 5	Venom or Salivary	Polistes Pol 5	205	P35780.1	549188	7
Polistes gallicus	Paper wasp	Pol g 5	Venom or Salivary	Polistes Pol 5	206	P83377.1	25091511	7
Polistes gallicus	Paper wasp	Unassigned	Venom or Salivary	Polistes Pol a 1 Pol d 1	42	P83542.1	41017429	7
Polybia paulista	wasp	Pol p 1.0101	Venom or Salivary	Polybia Pol p 1.0101 phospholipase	322	A2VBC4.1	166216292	9
Polybia paulista	wasp	Unassigned	Venom or Salivary	Polybia Pol p 1.0101 phospholipase	302	ADT89774.1	315190620	12
Polybia paulista	wasp	Unassigned	Venom or Salivary	Polybia Poly p 2 hyaluronidase	345	ADL09135.1	302201583	12
Polybia paulista	wasp	Poly p 2.0101	Venom or Salivary	Polybia Poly p 2 hyaluronidase	288	P86687.1	302425085	12
Polybia paulista	wasp	Unassigned	Venom or Salivary	Polybia Poly p 5, Poly s 5 venom allergen	141	ADD63684.1	290792375	11
Polybia paulista	wasp	Poly p 5.0102	Venom or Salivary	Polybia Poly p 5, Poly s 5 venom allergen	207	P86686.1	302595972	12
Polybia scutellaris rioplatensis	Wasp	Unassigned	Venom or Salivary	Polybia Poly p 5, Poly s 5 venom allergen	207	Q7Z156.2	47117356	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Pontastacus leptodactylus	Danube crayfish	Pon l 4.0101	Food Animal	Pontastacus Pon l 4	192	P05946.1	134309	15
Portunus pelagicus	Blue swimmer crab	Por p 1.0101	Food Animal	Portunus Por p 1 tropomyosin	284	AGE44125.1	448278534	14
Portunus sanguinolentus	Crab	Unassigned	Food Animal	Portunus Por p 1.0101 tropomyosin	284	ABL89183.1	119674937	8
Portunus trituberculatus	Crab	Unassigned	Food Animal	Portunus Por p 1.0101 tropomyosin	284	ABS12234.1	151505281	9
Procambarus clarkii	Red swamp crayfish	Unassigned	Food Animal	Procambarus red crayfish arginine kinase	357	AFA45339.1	375298901	13
Procambarus clarkii	Red swamp crayfish	Unassigned	Food Animal	Procambarus tropomysin	284	ACN87223.1	225348412	10
Prosopis juliflora	mesquite	Unassigned	Unassigned	Prosopis juliflora Pro j 1.0101	150	AKV72167.1	914410008	16
Prosopis juliflora	mesquite	Pro j 2.0101	Aero Plant	Prosopis Pro j 2	133	AHY24177.1	625293889	15
Protortonia cacti	Arthropod	Unassigned	Food Animal	Protortonia	335	BAH59276.1	237769615	11
Prunus armeniaca	Apricot	Pru ar 1	Food Plant	Prunus PRP (Bet v 1 family)	160	AAB97141.1	2677826	7
Prunus armeniaca	Apricot	Unassigned	Food Plant	Prunus Pru 3	119	ADR66945.1	313575730	12
Prunus armeniaca	Apricot	Unassigned	Food Plant	Prunus Pru 3	117	ADR66946.1	313575732	12
Prunus armeniaca	Apricot	Pru ar 3.0101	Food Plant	Prunus Pru 3	117	ADR66947.1	313575734	12
Prunus armeniaca	Apricot	Unassigned	Food Plant	Prunus Pru 3	117	ADR66948.1	313575736	12
Prunus avium	Cherry	Pru av 1.0101	Food Plant	Prunus PRP (Bet v 1 family)	160	AAC02632.1	1513216	7
Prunus avium	Cherry	Pru av 1.0203	Food Plant	Prunus PRP (Bet v 1 family)	160	AAS47037.1	44409496	7
Prunus avium	Cherry	Pru av 1.0202	Food Plant	Prunus PRP (Bet v 1 family)	160	AAS47036.1	44409474	7
Prunus avium	Cherry	Pru av 1.0201	Food Plant	Prunus PRP (Bet v 1 family)	160	AAS47035.1	44409451	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Prunus avium	Cherry	Unassigned	Food Plant	Prunus PRP (Bet v 1 family)	159	1H2O_A	159162378	9
Prunus avium	Cherry	Pru av 3	Food Plant	Prunus Pru 3	117	AAF26449.1	6715520	7
Prunus avium	Cherry	Unassigned	Food Plant	Prunus Pru 3	117	ADR66943.1	313575726	12
Prunus avium	Cherry	Unassigned	Food Plant	Prunus Pru 3	117	ADR66944.1	313575728	12
Prunus avium	Cherry	Pru av 4	Food Plant	Prunus Pru 4 Profilin peach cherry almond	131	AAD29411.1	4761582	7
Prunus avium	Cherry	Pru av 2	Food Plant	Prunus Pru av 2	245	AAB38064.1	1144346	7
Prunus domestica	Plum	Pru d 3	Food Plant	Prunus Pru 3	91	P82534.1	9297015	7
Prunus dulcis	Almond	Unassigned	Food Plant	Prunus persica Pru p 2 IUIS	241	ACE80974.1	190613941	10
Prunus dulcis	Almond	Pru du 4.0101	Food Plant	Prunus Pru 4 Profilin peach cherry almond	131	AAL91662.1	24473794	7
Prunus dulcis	Almond	Unassigned	Food Plant	Prunus Pru du 6 Amandin	531	3EHK_A	258588247	11
Prunus dulcis	Almond	Unassigned	Food Plant	Prunus Pru du 6 Amandin	178	AGR27935.1	523916668	15
Prunus dulcis	Almond	Pru du 6.0101	Food Plant	Prunus Pru du 6 Amandin	551	ADN39440.1	307159112	15
Prunus dulcis	Almond	Pru du 6.0201	Food Plant	Prunus Pru du 6 Amandin	504	ADN39441.1	307159114	15
Prunus dulcis	Almond	Unassigned	Food Plant	Prunus Seed allergenic protein 2 (Conglutin gamma)	25	P82952.1	75107131	8
Prunus dulcis x Prunus persica	Plant hybrid	Unassigned	Food Plant	Prunus persica Pru p 2 IUIS	160	ACE80939.1	190613871	10
Prunus dulcis x Prunus persica	Plant hybrid	Unassigned	Food Plant	Prunus persica Pru p 2 IUIS	246	ACE80956.1	190613905	10
Prunus dulcis x Prunus persica	Plant hybrid	Unassigned	Food Plant	Prunus persica Pru p 2 IUIS	246	ACE80958.1	190613909	10

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Prunus dulcis x Prunus persica	Plant hybrid	Pru p 2.0201	Food Plant	Prunus persica Pru p 2 IUIS	246	ACE80957.1	190613907	10
Prunus dulcis x Prunus persica	Plant hybrid	Pru p 2.0101	Food Plant	Prunus persica Pru p 2 IUIS	246	ACE80959.1	190613911	10
Prunus dulcis x Prunus persica	Plant hybrid	Pru p 2.0301	Food Plant	Prunus persica Pru p 2 IUIS	242	ACE80955.1	190613903	10
Prunus dulcis x Prunus persica	Plant hybrid	Unassigned	Food Plant	Prunus Pru 4 Profilin peach cherry almond	131	ACE80972.1	190613937	10
Prunus persica	Peach	Unassigned	Food Plant	Prunus persica Pru p 2 IUIS	246	P83332.1	25091405	12
Prunus persica	Peach	Unassigned	Food Plant	Prunus persica Pru p 2 IUIS	242	P83335.1	25091406	12
Prunus persica	Peach	Unassigned	Food Plant	Prunus persica Pru p 2 IUIS	246	AEV57471.1	359744030	13
Prunus persica	Peach	Pru p 1.0101	Food Plant	Prunus PRP (Bet v 1 family)	160	ABB78006.1	82492265	7
Prunus persica	Peach	Unassigned	Food Plant	Prunus PRP (Bet v 1 family)	160	AJE61291.1	748758672	16
Prunus persica	Peach	Unassigned	Food Plant	Prunus PRP (Bet v 1 family)	160	AJE61290.1	748758670	16
Prunus persica	Peach	Pru p 3.0101	Food Plant	Prunus Pru 3	91	P81402.1	3287877	7
Prunus persica	Peach	Unassigned	Food Plant	Prunus Pru 3	117	AAV40850.1	54793477	7
Prunus persica	Peach	Unassigned	Food Plant	Prunus Pru 3	117	ADR66939.1	313575718	12
Prunus persica	Peach	Unassigned	Food Plant	Prunus Pru 3	117	AGW21344.1	544369592	15
Prunus persica	Peach	Pru p 4.01	Food Plant	Prunus Pru 4 Profilin peach cherry almond	131	CAD37201.1	27528310	7
Prunus persica	Peach	Pru p 4.02	Food Plant	Prunus Pru 4 Profilin peach cherry almond	131	CAD37202.1	27528312	7
Prunus persica	Peach	Pru p 7.0101	Food Plant	Prunus Pru p 7 Pru m 7 Peamaclein	63	P86888.1	408407790	14

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Pseudocardium sachalinensis	Mollusc	Unassigned	Food Animal	Pseudocardium tropomyosin	284	BAH10154.1	219806598	10
Punica granatum	Pomegranate	Unassigned	Unassigned	Punica peptide Pommaclein	20	C0HKC0.1	1179881437	18
Punica granatum	Pomegranate	Pun g 1.0101	Food Plant	Punica Pun g 1	120	AHB19227.1	559797767	15
Punica granatum	Pomegranate	Pun g 1.0201	Food Plant	Punica Pun g 1	120	AHB19226.1	559797765	15
Punica granatum	Pomegranate	Pun g 1.0301	Food Plant	Punica Pun g 1	120	AHB19225.1	559797763	15
Pyrus communis	Pear	Pyr c 3.0101	Food Plant	Pyrus LTP Pyr c 3 IUIS	115	AAF26451.1	6715524	11
Pyrus communis	Pear	Unassigned	Food Plant	Pyrus LTP Pyr c 3 IUIS	94	AET05733.1	355525862	13
Pyrus communis	Pear	Unassigned	Food Plant	Pyrus LTP Pyr c 3 IUIS	94	AET05732.1	355525860	13
Pyrus communis	Pear	Unassigned	Food Plant	Pyrus LTP Pyr c 3 IUIS	94	AET05730.1	355525856	13
Pyrus communis	Pear	Pyr c 1.0101	Food Plant	Pyrus Pyr c 1	159	O65200.1	14423877	9
Pyrus communis	Pear	Pyr c 4	Food Plant	Pyrus Pyr c 4	131	AAD29410.1	4761580	7
Pyrus communis	Pear	Pyr c 5	Food Plant	Pyrus Pyr c 5	308	AAC24001.1	3243234	7
Quercus alba	Oak	Que a 1.0201	Aero Plant	Quercus Que a 1	159	ABZ81045.1	167472847	10
Quercus alba	Oak	Que a 1.0401	Aero Plant	Quercus Que a 1	160	ABZ81047.1	167472851	10
Quercus alba	Oak	Que a 1.0301	Aero Plant	Quercus Que a 1	160	ABZ81046.1	167472849	10
Rana esculenta	Frog	Ran e 1	Food Animal	Rana Ran e 1	110	CAC83046.1	20796729	7
Rana esculenta	Frog	Ran e 2	Food Animal	Rana Ran e 2	109	CAC95152.1	20797081	7
Rana sp. CH-2001	Frog	Unassigned	Food Animal	Rana Ran e 1	110	CAC83047.1	20796733	7
Rana sp. CH-2001	Frog	Unassigned	Food Animal	Rana Ran e 2	109	CAC95153.1	20797085	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Rattus norvegicus	Rat	Rat n 1	Aero Animal	Rattus Rat n 1	181	P02761.1	127533	7
Rattus norvegicus	Rat	Rat n 1	Aero Animal	Rattus Rat n 1	181	Q63213	81890324	7
Rattus norvegicus	Rat	Rat n 1.0101	Aero Animal	Rattus Rat n 1	177	AAA41198.1	204261	15
Rhizopus oryzae	Fungus	Unassigned	Aero Fungi	Rhizopus Rhi o 1.0101	401	AIS82657.1	695094784	16
Rhodotorula mucilaginosa	Fungus	Rho m 1.0101	Aero Fungi	Rhodotorula Rho m 1	439	AAP30720.1	30314940	11
Rhodotorula mucilaginosa	Fungus	Rho m 2.0101	Aero Fungi	Rhodotorula Rho m 2	342	AAT37679.1	54654335	7
Ricinus communis	Castor bean	Ric c 1.0101	Food Plant	Ricinus Ric c 1	258	CAA38097.1	21068	15
Rubus idaeus	raspberry	Rub i 1.0101	Food Plant	Rubus Rub i 1	137	ABG54495.1	110180525	8
Rubus idaeus	raspberry	Rub i 3.0101	Food Plant	Rubus Rub i 3	117	ABG54494.1	110180523	8
Salmo salar	Salmon	Sal s 1	Food Animal	Salmo Sal s 1	108	Q91483.3	18281421	7
Salmo salar	Salmon	Unassigned	Food Animal	Salmo Sal s 1	109	ACI68103.1	209734468	10
Salmo salar	Salmon	Sal s 1.0101	Food Animal	Salmo Sal s 1	109	CAA66403.1	1322183	15
Salmo salar	Salmon	Unassigned	Food Animal	Salmo Sal s 2 enolase	432	CBL79146.1	385145180	13
Salmo salar	Salmon	Sal s 2.0101	Food Animal	Salmo Sal s 2 enolase	434	ACH70931.1	197632415	15
Salmo salar	Salmon	Unassigned	Food Animal	Salmo Sal s 3 aldolase	363	CBL79147.1	385145176	13
Salmo salar	Salmon	Sal s 3.0101	Food Animal	Salmo Sal s 3 aldolase	363	NP_001133181.1	213511774	15
Salsola kali	Thistle	Unassigned	Unassigned	Salsola kali Sal k 6.01	381	AHL24657.1	589912883	15
Salsola kali	Thistle	Sal k 6	Aero Plant	Salsola kali Sal k 6.01	401	ARS33724.1	1194995727	18
Salsola kali	Thistle	Sal k 1.0201	Aero Plant	Salsola pectin methylesterase Sal k 1.01	362	AAT99258.1	51242679	8
Salsola kali	Thistle	Sal k 1.0302	Aero Plant	Salsola pectin methylesterase Sal k 1.01	339	AAX11261.1	59895728	8
Salsola kali	Thistle	Sal k 1.0301	Aero Plant	Salsola pectin methylesterase Sal k 1.01	339	AAX11262.1	59895730	8
Salsola kali	Thistle	Unassigned	Aero Plant	Salsola pectin methylesterase Sal k 1.01	339	ACO34813.1	225810597	10
Salsola kali	Thistle	Sal k 1.0101	Aero Plant	Salsola Sal k 1	42	P83181.1	25090947	10

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Salsola kali	Thistle	Unassigned	Aero Plant	Salsola Sal k 3 pollen allergen	757	ACO34814.1	225810599	10
Salsola kali	Thistle	Sal k 4.0101	Aero Plant	Salsola Sal k 4 profilin	133	ACS34771.1	239916566	11
Salsola kali	Thistle	Unassigned	Aero Plant	Salsola Sal k 4 profilin	133	AHL24658.1	589912885	15
Salsola kali	Thistle	Sal k 4.0201	Aero Plant	Salsola Sal k 4 profilin	133	ADK22841.1	300490499	15
Salsola kali	Thistle	Sal k 5.0101	Aero Plant	Salsola Sal k 5	151	ADK22842.1	300490501	15
Salvelinus fontinalis	Brook trout	Unassigned	Food Animal	Salvelinus parvalbumin	109	CAX32966.1	288557438	11
Salvelinus fontinalis	Brook trout	Unassigned	Food Animal	Salvelinus parvalbumin	108	CAX32967.1	288557440	11
Sarcoptes scabiei	mite	Unassigned	Unassigned	Sarcoptes scabiei paramyosin	828	SHD75397.1	1109598142	18
Sarcoptes scabiei type hominis	Scabies mite	Unassigned	Venom or Salivary	Sarcoptes Apolipoprotein Ssag1.2	330	AAO15613.1	27462848	7
Sarcoptes scabiei type hominis	Scabies mite	Unassigned	Venom or Salivary	Sarcoptes cysteine protease CO8	340	AAS93669.1	46406002	7
Sarcoptes scabiei type hominis	Scabies mite	Unassigned	Venom or Salivary	Sarcoptes cysteine proteases FO4	338	AAS93674.1	46406012	7
Sarcoptes scabiei type hominis	Scabies mite	Unassigned	Venom or Salivary	Sarcoptes cysteine proteases FO4	339	AAS93675.1	46406014	7
Sarcoptes scabiei type hominis	Scabies mite	Unassigned	Venom or Salivary	Sarcoptes cysteine proteases FO4	273	AAS93676.1	46406016	7
Sarcoptes scabiei type hominis	Scabies mite	Unassigned	Venom or Salivary	Sarcoptes Glutathione S-transferase Mu	219	AAO15607.1	27462836	7
Sarcoptes scabiei type hominis	Scabies mite	Unassigned	Venom or Salivary	Sarcoptes Glutathione S-transferase Mu	219	AAX37321.1	60920770	7
Sarcoptes scabiei type suis	Scabies mite	Unassigned	Aero Mite	Sarcoptes Apolipoprotein Ssag1.2	310	AGM48615.1	507480520	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Sardinops sagax	South American pilchard	Sar sa 1.0101	Food Animal	Sardinops Sar sa 1 parvalbumin	109	CAQ68366.1	193247972	10
Scapharca broughtonii	Clam	Unassigned	Food Animal	Scapharca tropomyosin	284	BAH10151.1	219806592	10
Schedonorus arundinaceus	Tall fescue	Unassigned	Aero Plant	Festuca group 1 allergen	35	Q7M1Y1	75139991	7
Schedonorus arundinaceus	Tall fescue	Unassigned	Aero Plant	Festuca group 1 allergen	17	C37396	320610	7
Schedonorus arundinaceus	Tall fescue	Unassigned	Aero Plant	Festuca group 1 allergen	20	D37396	320611	7
Schistosoma japonicum	Schistosoma	Unassigned	Protozoan	Schistosoma profilin	129	AAP06493.1	29841461	7
Schistosoma japonicum	Schistosoma	Unassigned	Protozoan	Schistosoma tegumental antigen	191	AAC67308.1	2739154	7
Schizophyllum commune H4-8	Mushroom	Sch c 1.0101	Food Fungi	Schizophyllum Sch c 1	576	XP_003030591.1	302681819	15
Scleronephthya gracillimum		Unassigned	Unassigned	Scleronephthya	225	BAW32538.1	1113818793	18
Scleronephthya gracillimum		Unassigned	Unassigned	Scleronephthya	225	BAW32537.1	1113818791	18
Scleronephthya gracillimum		Unassigned	Unassigned	Scleronephthya	225	BAW32536.1	1113818789	18
Scleronephthya gracillimum		Unassigned	Unassigned	Scleronephthya	225	BAW32535.1	1113818787	18
Scomber japonicus	Chub mackerel	Unassigned	Food Animal	Scomber Parvalbumin	109	BAC66618.1	29420793	7
Scomber scombrus	Atlantic mackerel	Unassigned	Food Animal	Scomber Parvalbumin	109	CAX32965.1	288557436	11
Scylla paramamosain	green mud crab	Unassigned	Food Animal	Scylla arginine kinase	357	AFA45340.1	375298903	13
Scylla paramamosain	green mud crab	Unassigned	Unassigned	Scylla paramamosain Sarc Ca Binding Ptn	193	AFJ80778.1	387571563	18
Scylla serrata	giant mud crab	Unassigned	Food Animal	Scylla sp. (mud crab) tropomyosin	284	ABS12233.1	151505279	9
Sebastes marinus	ocean perch (red fish)	Seb m 1.0101	Food Animal	Sebastes Seb m 1	109	CAQ72968.1	242253959	11
Sebastes marinus	ocean perch (red fish)	Seb m 1.0201	Food Animal	Sebastes Seb m 1	110	CAQ72969.1	242253961	11

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Secale cereale	Rye	Sec c 20.0101	Food Plant	Secale Sec c 20	23	AAB37403.1	1699225	15
Secale cereale	Rye	Sec c 20.0201	Food Plant	Secale Sec c 20	29	AAB37406.1	1699228	15
Secale cereale	Rye	Sec c 38.0101	Food Plant	Secale Sec c 38.01	26	AAB34365.1	994865	10
Secale cereale	Rye	Unassigned	Aero Plant	Secale Sec c 4	520	CAH92630.1	55859456	7
Secale cereale	Rye	Unassigned	Aero Plant	Secale Sec c 4	518	CAH92627.1	55859454	7
Secale cereale	Rye	Unassigned	Aero Plant	Secale Sec c 5	16	Q7M263	75140047	7
Secale cereale	Rye	Sec c 5.0101	Food Plant	Secale Sec c 5	292	CBG76811.1	332205751	12
Sepia esculenta	Cuttlefish	Unassigned	Food Animal	Sepia tropomyosin	284	BAE54429.1	83715928	7
Sepioteuthis lessoniana	Bigfin reef squid	Unassigned	Food Animal	Sepioteuthis tropomyosin	284	BAE54430.1	83715930	7
Sesamum indicum	Sesame	Unassigned	Food Plant	Sesamum seed maturation-like protein	345	ACB55491.1	171853012	16
Sesamum indicum	Sesame	Ses i 1	Food Plant	Sesamum Ses i 1	153	AAK15088.1	13183175	7
Sesamum indicum	Sesame	Unassigned	Food Plant	Sesamum Ses i 1	153	ACI41244.1	209165427	10
Sesamum indicum	Sesame	Ses i 2	Food Plant	Sesamum Ses i 2	148	AAD42943.1	5381323	7
Sesamum indicum	Sesame	Ses i 3	Food Plant	Sesamum Ses i 3	585	AAK15089.1	13183177	7
Sesamum indicum	Sesame	Ses i 4.0101	Food Plant	Sesamum Ses i 4 oleosin	166	AAG23840.1	10834827	13
Sesamum indicum	Sesame	Unassigned	Food Plant	Sesamum Ses i 5 oleosin	145	ACH85188.1	198250343	10
Sesamum indicum	Sesame	Ses i 5.0101	Food Plant	Sesamum Ses i 5 oleosin	145	AAD42942.1	5381321	15
Sesamum indicum	Sesame	Ses i 6.0101	Food Plant	Sesamum Ses i 6	459	AAD42944.1	5381325	15
Sesamum indicum	Sesame	Ses i 7.0101	Food Plant	Sesamum Ses i 7	497	AAK15087.1	13183173	15
Sinapis alba	White mustard	Sin a 1	Food Plant	Sinapis Sin a 1.01	145	CAA62909.1	1009434	7
Sinapis alba	White mustard	Sin a 1	Food Plant	Sinapis Sin a 1.01	145	CAA62910.1	1009436	7
Sinapis alba	White mustard	Sin a 1	Food Plant	Sinapis Sin a 1.01	145	CAA62911.1	1009438	7
Sinapis alba	White mustard	Sin a 1	Food Plant	Sinapis Sin a 1.01	145	CAA62912.1	1009440	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Sinapis alba	White mustard	Sin a 1	Food Plant	Sinapis Sin a 1.01	145	CAA62908.1	1009442	7
Sinapis alba	White mustard	Sin a 1.0101	Food Plant	Sinapis Sin a 1.01	145	P15322.2	51338758	7
Sinapis alba	White mustard	Sin a 2.0101	Food Plant	Sinapis Sin a 2.01 11S globulin	510	AAX77383.1	62240390	7
Sinapis alba	White mustard	Unassigned	Food Plant	Sinapis Sin a 2.01 11S globulin	523	AAX77384.1	62240392	7
Sinapis alba	White mustard	Sin a 3.0101	Food Plant	Sinapis Sin a 3.01 LTP	92	ABU95411.1	156778059	12
Sinapis alba	White mustard	Sin a 4.0101	Food Plant	Sinapis Sin a 4.01 profilin	131	ABU95412.1	156778061	12
Sinonovacula constricta	Chinese razor clam	Unassigned	Food Animal	Sinonovacula tropomyosin [Song paper]	284	ABU53681.1	156145810	15
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Unassigned	Unassigned	Solanum lycopersicum Sola l 6	96	NP_001306883.1	985801667	17
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Unassigned	Unassigned	Solanum lycopersicum Sola l 7	115	NP_001316123.1	1042161070	17
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Sola l 1.0101	Food Plant	Solanum Sola l 1 profilin (Lyc e 1)	131	CAD10377.1	16555787	7
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Lyc e 1	Food Plant	Solanum Sola l 1 profilin (Lyc e 1)	131	AAL29690.1	17224229	7
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Sola l 2.0101	Food Plant	Solanum Sola l 2 Beta-fructofuranosidase (Lyc e 2)	553	AAL75449.1	18542113	7
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Sola l 2.0201	Food Plant	Solanum Sola l 2 Beta-fructofuranosidase (Lyc e 2)	636	AAL75450.1	18542115	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Unassigned	Food Plant	Solanum Sola l 3 LTP (Lyc e 3)	114	CAJ19705.1	71360928	7
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Sola l 3.0101	Food Plant	Solanum Sola l 3 LTP (Lyc e 3)	114	AAB42069.1	1816535	15
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Sola l 4.0101	Food Plant	Solanum Sola l 4 PR-10 (Lyc e 4)	178	CAA75803.1	2887310	14
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Sola l 4.0201	Food Plant	Solanum Sola l 4 PR-10 (Lyc e 4)	160	AHC08074.1	565380268	15
Solanum lycopersicum (Lycopersicon esculentum)	Tomato	Unassigned	Food Plant	Solanum Sola l 4 PR-10 (Lyc e 4)	160	AHC08073.1	565380238	15
Solanum tuberosum	Potato	Unassigned	Food Plant	Solanum profilin-like	131	ABA81885.1	77416979	7
Solanum tuberosum	Potato	Unassigned	Food Plant	Solanum profilin-like	131	ABB16985.1	77999277	7
Solanum tuberosum	Potato	Unassigned	Food Plant	Solanum Sola t 1	386	CAA31575.1	21510	7
Solanum tuberosum	Potato	Unassigned	Food Plant	Solanum Sola t 1	386	CAA27571.1	21512	7
Solanum tuberosum	Potato	Unassigned	Food Plant	Solanum Sola t 1	386	CAA27588.1	21514	7
Solanum tuberosum	Potato	Unassigned	Food Plant	Solanum Sola t 1	386	AAA33819.1	169500	7
Solanum tuberosum	Potato	Sola t 1	Food Plant	Solanum Sola t 1	386	P15476.2	158517845	9
Solanum tuberosum	Potato	Sola t 2	Food Plant	Solanum Sola t 2	188	P16348.1	124148	7
Solanum tuberosum	Potato	Sola t 3	Food Plant	Solanum Sola t 3	222	P20347.3	20141344	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Solanum tuberosum	Potato	Sola t 3.0101	Food Plant	Solanum Sola t 3	186	AAB63099.1	1575306	15
Solanum tuberosum	Potato	Sola t 4.0101	Food Plant	Solanum Sola t 4	221	BAA04149.1	994779	15
Solen strictus	Gould's razor shell	Unassigned	Food Animal	Solen tropomyosin	284	BAH10156.1	219806602	10
Solenopsis geminata	Tropical Fire Ant	Sol g 4.0101	Venom or Salivary	Solenopsis Sol g 4 Sol i 4	137	AAF65312.1	7638028	7
Solenopsis geminata	Tropical Fire Ant	Sol g 4.0201	Venom or Salivary	Solenopsis Sol g 4 Sol i 4	137	AAF65313.1	7638030	7
Solenopsis invicta	Red fire ant	Sol i 4	Venom or Salivary	Solenopsis Sol g 4 Sol i 4	137	AAC97370.1	4038411	7
Solenopsis invicta	Red fire ant	Sol i 4.0101	Venom or Salivary	Solenopsis Sol g 4 Sol i 4	137	AAC97369.1	4038409	11
Solenopsis invicta	Red fire ant	Unassigned	Venom or Salivary	Solenopsis Sol i 1	58	AAB36117.1	1336809	7
Solenopsis invicta	Red fire ant	Unassigned	Venom or Salivary	Solenopsis Sol i 1	25	AAB36119.1	1336811	7
Solenopsis invicta	Red fire ant	Unassigned	Venom or Salivary	Solenopsis Sol i 1	26	AAB36120.1	1336812	7
Solenopsis invicta	Red fire ant	Unassigned	Venom or Salivary	Solenopsis Sol i 1	26	AAB36121.1	1336813	7
Solenopsis invicta	Red fire ant	Sol i 1.0101	Venom or Salivary	Solenopsis Sol i 1	346	AAT95008.1	51093373	7
Solenopsis invicta	Red fire ant	Sol i 2.0101	Venom or Salivary	Solenopsis Sol i and Sol r Venom allergen II	138	P35775.1	549179	7
Solenopsis invicta	Red fire ant	Sol i 3.0101	Venom or Salivary	Solenopsis Venom allergen III	234	AAB65434.1	2293571	11
Solenopsis richteri	Black fire ant	Sol r 2.0101	Venom or Salivary	Solenopsis Sol i and Sol r Venom allergen II	119	P35776.2	6136162	7
Solenopsis richteri	Black fire ant	Sol r 3.0101	Venom or Salivary	Solenopsis Venom allergen III	211	P35779.2	6136163	7
Solenopsis saevissima	Brazilian fire ant	Unassigned	Venom or Salivary	Solenopsis Sol g 4 Sol i 4	137	ADD74392.1	291092710	12
Sorghum halepense	Johnson grass	Sor h 2.0201	Unassigned	Sorghum halepens group 2 allergen	121	AIL01319.1	674275735	16
Sorghum halepense	Johnson grass	Sor h 2.0101	Unassigned	Sorghum halepens group 2 allergen	119	AIL01318.1	674275733	16

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Sorghum halepense	Johnson grass	Sor h 1.0101	Aero Plant	Sorghum Sor h 1	266	AIL01316.1	674275729	15
Sorghum halepense	Johnson grass	Sor h 1.0201	Aero Plant	Sorghum Sor h 1	266	AIL01317.1	674275731	15
Sorghum halepense	Johnson grass	Sor h 13.0101	Aero Plant	Sorghum Sor h 13	422	AIL01320.1	674275737	15
Sorghum halepense	Johnson grass	Sor h 13.0201	Aero Plant	Sorghum Sor h 13	410	AIL01321.1	674275739	15
Stachybotrys chartarum	Fungus	Sta 3.0101	Aero Fungi	Stachybotrys Sta c 3	144	ACT37324.1	253970748	14
Staphylococcus aureus	Bacteria	Unassigned	Bacteria skin	Staphylococcus enterotoxin SEA	233	1ESF_B	1633233	9
Staphylococcus aureus	Bacteria	Unassigned	Bacteria skin	Staphylococcus enterotoxin SEB	254	CAJ43561.1	83308249	9
Staphylococcus aureus	Bacteria	Unassigned	Bacteria skin	Staphylococcus enterotoxin SEC	266	P34071.1	462026	9
Staphylococcus aureus	Bacteria	Unassigned	Bacteria skin	Staphylococcus enterotoxin SED	258	P20723.1	119654	9
Staphylococcus aureus	Bacteria	Unassigned	Bacteria skin	Staphylococcus enterotoxin TSST 1	234	P06886.1	136457	9
Stemphylium callistephi	Fungus	Unassigned	Aero Fungi	Stemphylium major allergen alt a1-like	137	AAT66567.1	49476467	7
Stemphylium sp. CID1012	Fungus	Unassigned	Aero Fungi	Stemphylium major allergen alt a1-like	137	ABS29033.1	152060760	9
Stemphylium vesicarium	Fungus	Unassigned	Aero Fungi	Stemphylium major allergen alt a1-like	137	AAT66566.1	49476465	7
Strongyloides stercoralis	Parasitic nematode	Unassigned	Worm (parasite)	Strongyloides L3NieAg.01	229	AAD46493.1	5669875	7
Suidasia medanensis	Mite	Unassigned	Aero Mite	Suidasia putative Sui m 2	141	AAS75831.1	45738062	7
Sus scrofa	Pig	Unassigned	Aero Animal	Sus Porcine Pepsin	385	P00791.3	118572685	11
Sus scrofa	Pig	Unassigned	Unassigned	Sus s serum albumin	605	AAA30988.1	164318	17
Sus scrofa	Pig	Unassigned	Unassigned	Sus s serum albumin	607	NP_001005208.1	52353352	17
Syringa vulgaris	Lilac	Syr v 3.0101	Aero Plant	Syringa Syr v 3	81	P58171.1	14423847	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Syringa vulgaris	Lilac	Syr v 1.0101	Aero Plant	Syringa Syr v I	145	S43242	631911	7
Syringa vulgaris	Lilac	Syr v 1.0102	Aero Plant	Syringa Syr v I	145	S43243	631912	7
Syringa vulgaris	Lilac	Syr v 1.0103	Aero Plant	Syringa Syr v I	145	S43244	631913	7
Tabanus yao	Horse Fly	Tab y 1.0101	Venom or Salivary	Tabanus Tab y 1 Apyrase	554	ADX78255.1	323473390	12
Tabanus yao	Horse Fly	Tab y 2.0101	Venom or Salivary	Tabanus Tab y 2 Hyaluronidase	349	ADM18346.1	304273371	12
Tabanus yao	Horse Fly	Tab y 5.0101	Venom or Salivary	Tabanus Tab y 5	256	ADM18345.1	304273369	12
Thaumetopoea pityocampa	Pine moth	Tha p 1.0101	Contact	Thaumetopoea Tha p 1 full length	126	ADK47876.1	301030229	12
Thaumetopoea pityocampa	Pine moth	Tha p 2.0101	Contact	Thaumetopoea Tha p 2	115	P86360.1	408387552	14
Thaumetopoea pityocampa	Pine moth	Unassigned	Unassigned	Thaumetopoea Tha p 2	104	CEE03319.1	1056731906	18
Thaumetopoea solitaria	Proces. moth	Unassigned	Unassigned	Thaumetopoea Tha p 2	100	CEE03318.1	1056731899	18
Theragra chalcogramma	Alaska pollock	Unassigned	Food Animal	Theragra parvalbumin	109	AAK63089.1	14531020	7
Theragra chalcogramma	Alaska pollock	Unassigned	Food Animal	Theragra parvalbumin	109	AAK63088.1	14531018	7
Thunnus albacares	Yellowfin tuna	Thu a 2.0101	Food Animal	Thunnus Thu a 2 enolase	432	CBL79145.1	385145178	13
Thunnus albacares	Yellowfin tuna	Unassigned	Food Animal	Thunnus Thu a 2 enolase	12	P86978.1	576011132	15
Thunnus albacares	Yellowfin tuna	Unassigned	Food Animal	Thunnus Thu a 3 aldolase	364	CAX62602.1	291195949	12
Thunnus albacares	Yellowfin tuna	Thu a 3.0101	Food Animal	Thunnus Thu a 3 aldolase	37	P86979.1	576011088	15
Todarodes pacificus	Japanese flying squid	Unassigned	Food Animal	Todarodes Tod p 1	284	BAE54431.1	83715932	7
Trachurus japonicus	Japanese horse mackerel	Unassigned	Food Animal	Trachurus parvalbumin	107	BAE46763.1	77799800	7
Tresus keenae	clam	Unassigned	Food Animal	Tresus tropomyosin	284	BAH10155.1	219806600	10

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Triatoma protracta	Western conenose	Tria p 1	Venom or Salivary	Triatoma Tria p 1	169	AAF07903.2	15426413	7
Trichophyton rubrum	Fungus	Tri r 2	Contact	Trichophyton (Arthroderma) Tri r 2	412	AAD52013.1	5813790	7
Trichophyton rubrum	Fungus	Tri r 4	Contact	Trichophyton tri 4 allergen (Arthroderma)	726	AAD52012.1	5813788	7
Trichophyton schoenleinii	Fungus	Unassigned	Contact	Trichophyton (Arthroderma) Tri r 2	405	Q8J077.1	74663809	12
Trichophyton schoenleinii	Fungus	Unassigned	Contact	Trichophyton tri 4 allergen (Arthroderma)	726	CAD23374.1	23894227	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 14 LTP_amylase inhibitor	113	P24296.2	417370	11
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum aestivum Tri a 40	143	CAA42453.1	21711	7
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum aestivum Tri a 40	143	ACG59281.1	195957140	10
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum aestivum Tri a 41	60	AKJ77988.1	827354845	16
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum aestivum Tri a 42	76	AKJ77986.1	827354790	16
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum aestivum Tri a 43	108	AKJ77987.1	827354822	16
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum aestivum Tri a 44	94	CAI64398.1	66840998	7
Triticum aestivum	Wheat	Tri a 44.0101	Aero Plant	Triticum aestivum Tri a 44	107	AKJ77990.1	827354912	16
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum aestivum Tri a 45	89	AKJ77985.1	827354784	16
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	307	CAA35238.1	21673	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum alpha/beta gliadin	286	CAA25593.1	21755	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	296	CAA26383.1	21757	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	286	CAA26384.1	21761	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	313	CAA26385.1	21765	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	318	AAA34275.1	170710	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	291	AAA34276.1	170712	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	313	AAA34279.1	170718	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	286	AAA34280.1	170720	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	262	AAA34281.1	170722	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	297	AAA34282.1	170724	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	282	AAA34283.1	170726	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	186	AAA34284.1	170728	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum alpha/beta gliadin	259	BAA12318.1	1304264	7
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum Bakers asthma allergen #4	27	P81496.1	3913017	7
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum flour Glutathione Transferase	222	ACE82289.1	190684057	11
Triticum aestivum	Wheat	Tri a 19.0101	Food Plant	Triticum omega-5 gliadin Tri a 19	439	BAE20328.1	73912496	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum omega-5 gliadin Tri a 19	359	CAR82265.1	208605344	10
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum omega-5 gliadin Tri a 19	272	CAR82266.1	208605346	10
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum omega-5 gliadin Tri a 19	346	CAR82267.1	208605348	10
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum omega-5 gliadin Tri a 19	366	BAN29067.1	508732623	15
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum putative leucine-rich repeat protein	137	CAI64397.1	66840996	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum serine carboxypeptidase II	260	CAI64396.1	66840994	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum serine carboxypeptidase II	444	P08819.2	125987805	10

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Thaumatin-like	173	P27357.1	135917	12
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum Tri a 12	131	ACE82291.1	190684061	11
Triticum aestivum	Wheat	Tri a 12.0103	Aero Plant	Triticum Tri a 12	131	CAA61945.2	548948852	14
Triticum aestivum	Wheat	Tri a 12.0101	Aero Plant	Triticum Tri a 12	131	CAA61943.2	548948848	15
Triticum aestivum	Wheat	Tri a 12.0102	Aero Plant	Triticum Tri a 12	131	CAA61944.2	548948850	15
Triticum aestivum	Wheat	Tri a 12.0104	Aero Plant	Triticum Tri a 12	131	CAQ57979.1	207366248	15
Triticum aestivum	Wheat	Tri a 15.0101	Gliadin	Triticum Tri a 15	121	CBA13560.1	283465829	11
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 20	302	AAA34272.1	170702	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 20	291	AAA34274.1	170708	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 20	251	AAA34288.1	170736	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 20	327	AAA34289.1	170738	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 20	279	BAA11251.1	1063270	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 20	285	CAI78902.1	62484809	7
Triticum aestivum	Wheat	Tri a 20.0101	Gliadin	Triticum Tri a 20	279	BAN29066.1	508732621	15
Triticum aestivum	Wheat	Tri a 21.0101	Gliadin	Triticum Tri a 21 alpha, beta-gliadin	281	CAY54134.1	283476402	11
Triticum aestivum	Wheat	Tri a 25.0101	Gliadin	Triticum Tri a 25	125	CAB96931.1	8980491	15
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 26	830	CAA43331.1	21743	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 26	648	CAA31396.1	21751	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 26	660	CAA26847.1	21779	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 26	39	CAA24934.1	21793	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 26	705	CAA43361.1	22090	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 26	815	AAB02788.1	170743	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 26	838	CAA27052.1	736319	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 26	101	CAA24933.1	897811	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 26	794	BAN29068.1	508732625	15
Triticum aestivum	Wheat	Tri a 26.0101	Gliadin	Triticum Tri a 26	848	CAA31395.4	288860106	15
Triticum aestivum	Wheat	Tri a 26.0201	Gliadin	Triticum Tri a 26	795	AAZ23584.1	71084277	15
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 27.0101 Thiol reductase	203	BAC76688.1	30793446	7
Triticum aestivum	Wheat	Tri a 28.0101	Gliadin	Triticum Tri a 28	119	CAI84642.1	66841026	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 29	145	CAA35598.1	21701	7
Triticum aestivum	Wheat	Tri a 29.0101	Aero Plant	Triticum Tri a 29	120	CAZ76052.1	253783731	11
Triticum aestivum	Wheat	Tri a 29.0201	Aero Plant	Triticum Tri a 29	120	CBA13559.1	283465827	11
Triticum aestivum	Wheat	Tri a 30.0101	Food Plant	Triticum Tri a 30	168	CAA35597.1	21713	7
Triticum aestivum	Wheat	Tri a 31.0101	Food Plant	Triticum Tri a 31	253	CAC14917.1	11124572	7
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum Tri a 32 Peroxiredoxin	218	ACE82290.1	190684059	11
Triticum aestivum	Wheat	Tri a 32.0101	Aero Plant	Triticum Tri a 32 Peroxiredoxin	218	Q6W8Q2.1	75324900	14
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 33 Serine protease inhibitor	399	CAA72273.1	1885350	7
Triticum aestivum	Wheat	Tri a 33.0101	Gliadin	Triticum Tri a 33 Serine protease inhibitor	398	CAB52710.1	5734506	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Triticum aestivum	Wheat	Tri a 34.0101	Gliadin	Triticum Tri a 34 GAPDH	337	CAZ76054.1	253783729	11
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 36	307	CAA31685.1	21773	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 36	356	CAA30570.1	21783	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 36	304	AAA34285.1	170730	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 36	323	AAA34286.1	170732	7
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 36	244	AAA34287.1	170734	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 36	373	O22116	75317968	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 36	229	CAA59338.1	886963	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 36	261	CAA59339.1	886965	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 36	276	CAA59340.1	886967	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 36	285	O22108	75219081	7
Triticum aestivum	Wheat	Unassigned	Food Plant	Triticum Tri a 36	326	CAI79052.1	62550933	7
Triticum aestivum	Wheat	Tri a 36.0101	Food Plant	Triticum Tri a 36	369	AEH31546.1	335331566	12
Triticum aestivum	Wheat	Unassigned	Gliadin	Triticum Tri a 36	283	BAN29069.1	508732627	15
Triticum aestivum	Wheat	Tri a 37.0101	Food Plant	Triticum Tri a 37 alpha purothionin	137	CAA65313.1	4007850	14
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum Tri a 39 serine proteinase inhibitor-lik	84	ABS58503.1	154101366	10
Triticum aestivum	Wheat	Unassigned	Aero Plant	Triticum Tri a 39 serine proteinase inhibitor-lik	84	P82977.2	122065237	11
Triticum aestivum	Wheat	Tri a 39.0101	Aero Plant	Triticum Tri a 39 serine proteinase inhibitor-lik	84	CCK33471.1	403213259	14

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Triticum monococcum subsp. aegilopoides	Wheat grass	Unassigned	Unassigned	Triticum alpha/beta gliadin	293	APY24042.1	1137166044	18
Triticum turgidum subsp. durum	Wheat	Unassigned	Food Plant	Triticum aestivum Tri a 40	143	CAA34709.1	21916	7
Triticum turgidum subsp. durum	Wheat	Unassigned	Food Plant	Triticum Tri a 29	145	CAA39099.1	21920	7
Triticum turgidum subsp. durum	Wheat	Unassigned	Food Plant	Triticum Tri a 36	295	CAA36063.1	21926	7
Triticum turgidum subsp. durum	Wheat	Unassigned	Food Plant	Triticum Tri a 36	285	CAA44473.1	21930	7
Triticum urartu	Wheat	Unassigned	Food Plant	Triticum alpha/beta gliadin	296	AAA34290.1	170740	7
Tyrophagus putrescentiae	Dust mite	Unassigned	Unassigned	Tyrophagus Blo-t-5-like loose group	135	AAX34057.1	60679590	9
Tyrophagus putrescentiae	Dust mite	Unassigned	Unassigned	Tyrophagus Blo-t-5-like loose group	128	AAX34058.1	60679592	9
Tyrophagus putrescentiae	Dust mite	Unassigned	Unassigned	Tyrophagus Blo-t-5-like loose group	138	AAX34059.1	60679594	9
Tyrophagus putrescentiae	Dust mite	Tyr p 28.0101	Unassigned	Tyrophagus putrescentiae Tyr p 28	659	AOD75395.1	1055365842	17
Tyrophagus putrescentiae	Dust mite	Tyr p 35.0101	Unassigned	Tyrophagus putrescentiae Tyr p 35	486	AOD75396.1	1055365860	17
Tyrophagus putrescentiae	Dust mite	Tyr p 36.0101	Unassigned	Tyrophagus putrescentiae Tyr p 36	131	AOD75399.1	1055365943	17
Tyrophagus putrescentiae	Dust mite	Unassigned	Aero Mite	Tyrophagus Tyr p 10 tropomyosin	284	ABQ96644.1	148615631	9
Tyrophagus putrescentiae	Dust mite	Unassigned	Aero Mite	Tyrophagus Tyr p 10 tropomyosin	201	ABU97479.1	156938915	9
Tyrophagus putrescentiae	Dust mite	Tyr p 10.0101	Aero Mite	Tyrophagus Tyr p 10 tropomyosin	284	AAT40866.1	48249227	9
Tyrophagus putrescentiae	Dust mite	Tyr p 13	Aero Mite	Tyrophagus Tyr p 13	131	AAU11502.1	51860756	7

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Tyrophagus putrescentiae	Dust mite	Unassigned	Aero Mite	Tyrophagus Tyr p 13	130	ABM53751.1	121296500	9
Tyrophagus putrescentiae	Dust mite	Unassigned	Aero Mite	Tyrophagus Tyr p 13	131	ABU97480.1	156938917	9
Tyrophagus putrescentiae	Dust mite	Tyr p 2	Aero Mite	Tyrophagus Tyr p 2	141	CAA73221.1	2182106	7
Tyrophagus putrescentiae	Dust mite	Tyr p 24.0101	Aero Mite	Tyrophagus Tyr p 24 Troponin C	153	ACL36923.1	219815476	11
Tyrophagus putrescentiae	Dust mite	Tyr p 3.0101	Aero Mite	Tyrophagus Tyr p 3	285	ABZ81991.1	167540622	11
Tyrophagus putrescentiae	Dust mite	Unassigned	Aero Mite	Tyrophagus Tyr p 8	218	AGG10560.1	452215228	14
Ulocladium alternariae	Fungus	Unassigned	Aero Fungi	Ulocladium alt a1-like	138	AAT66607.1	49476547	7
Ulocladium atrum	Fungus	Unassigned	Aero Fungi	Ulocladium alt a1-like	137	AAT66609.1	49476551	7
Ulocladium capsicum	Fungus pepper	Unassigned	Unassigned	Ulocladium alt a1-like	134	ACH42744.1	197110100	10
Ulocladium chartarum	Fungus	Unassigned	Aero Fungi	Ulocladium alt a1-like	137	AAT66610.1	49476553	7
Ulocladium dauci	Fungus	Unassigned	Unassigned	Ulocladium alt a1-like	137	ACJ65836.1	215399749	11
Ulocladium microsporum	Fungus	Unassigned	Unassigned	Ulocladium alt a1-like	138	AGC36415.1	441467668	18
Ulocladium oudemansii	Fungus	Unassigned	Unassigned	Ulocladium alt a1-like	137	ACH42743.1	197110098	10
Ulocladium oudemansii	Fungus	Unassigned	Unassigned	Ulocladium alt a1-like	138	ACI44002.1	209363467	10
Ulocladium sp. CID262	Fungus	Unassigned	Unassigned	Ulocladium alt a1-like	137	ABQ59259.1	148357923	9
Ulocladium sp. CID598	Fungus	Unassigned	Unassigned	Ulocladium alt a1-like	137	ABQ59258.1	148357921	9
Ulocladium sp. CID68	Fungus	Unassigned	Unassigned	Ulocladium alt a1-like	137	ABQ59255.1	148357915	9
Ulocladium sp. HSAUP1144	Fungus	Unassigned	Unassigned	Ulocladium alt a1-like	138	ACJ54737.1	213958825	11
Ulocladium sp. XGZ-2008	Fungus	Unassigned	Unassigned	Ulocladium alt a1-like	137	ACH42741.1	197110094	10

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Ulocladium sp. XGZ-2011a	Fungus	Unassigned	Unassigned	Ulocladium alt a1-like	137	AGC36416.1	441467671	18
Vachellia farnesiana	Acacia	Aca f 1	Unassigned	Acacia (Vachellia) Aca f 1	150	AKV72166.1	914410006	17
Vachellia farnesiana	Acacia	Aca f 2	Unassigned	Acacia (Vachellia) profilin Aca f 2	133	AIV43662.1	701225196	17
Venerupis philippinarum	Clam	Unassigned	Food Animal	Venerupis tropomyosin	284	BAH10157.1	219806573	10
Vespa affinis	Lesser banded hornet	Unassigned	Food Animal	Vespa affinis Phospholipase A1	334	P0DMB5.1	576011175	15
Vespa affinis	Lesser banded hornet	Unassigned	Food Animal	Vespa affinis Phospholipase A1	334	P0DMB4.1	576011171	15
Vespa crabro	European hornet	Unassigned	Venom or Salivary	Vespa Vesp c 1 phospholipase	301	P0CH87.1	313471397	12
Vespa crabro	European hornet	Vesp c 5.0101	Venom or Salivary	Vespa Vesp c 5	202	P35781.1	549184	7
Vespa crabro	European hornet	Vesp c 5.0102	Venom or Salivary	Vespa Vesp c 5	202	P35782.1	549185	7
Vespa magnifica	Hornet	Unassigned	Venom or Salivary	Vespa magnifica Vesp ma 2 hyaluronidase	357	CBY83816.1	315133295	12
Vespa magnifica	Hornet	Unassigned	Venom or Salivary	Vespa magnifica Vesp ma 5	225	CBY93636.1	319801357	12
Vespa mandarinia	Wasp	Vesp m 5.0101	Venom or Salivary	Vespa Vesp c 5	202	P81657.1	6136165	7
Vespula flavopilosa	Wasp	Ves f 5.0101	Venom or Salivary	Vespula Ves f 5	204	P35783.1	549189	7
Vespula germanica	Wasp	Unassigned	Venom or Salivary	Vespula Phospholipase A1 Ves m/v 1	300	CAJ28931.1	74035843	7
Vespula germanica	Wasp	Ves g 5.0101	Venom or Salivary	Vespula Ves f 5	204	P35784.1	549190	7
Vespula germanica	Wasp	Unassigned	Venom or Salivary	Vespula Ves f 5	204	CAJ28930.1	74035841	7
Vespula germanica	Wasp	Unassigned	Venom or Salivary	Vespula Ves v 2	331	CAL59818.1	116174180	8
Vespula germanica	Wasp	Unassigned	Venom or Salivary	Vespula Ves v 2	323	CAL59819.1	116174182	8
Vespula maculifrons	Wasp	Ves m 1.0101	Venom or Salivary	Vespula Phospholipase A1 Ves m/v 1	300	P51528.1	1709545	8

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Vespula maculifrons	Wasp	Ves m 5.0101	Venom or Salivary	Vespula Ves f 5	204	P35760.1	549191	7
Vespula maculifrons	Wasp	Unassigned	Venom or Salivary	Vespula Ves f 5	227	ABC73068.1	85681830	7
Vespula maculifrons	Wasp	Unassigned	Venom or Salivary	Vespula Ves m 2 Hyaluronidase	31	P0CH89.1	313118253	12
Vespula pensylvanica	Wasp	Ves p 5.0101	Venom or Salivary	Vespula Ves f 5	204	P35785.1	549192	7
Vespula squamosa	Wasp	Ves s 5.0101	Venom or Salivary	Vespula Ves f 5	205	P35786.1	549193	7
Vespula squamosa	Wasp	Unassigned	Venom or Salivary	Vespula Ves s 1 phospholipase	298	P0CH86.1	313471398	12
Vespula vidua	Wasp	Ves vi 5.0101	Venom or Salivary	Vespula Ves f 5	206	P35787.1	549194	7
Vespula vulgaris	Wasp	Ves v 1.0101	Venom or Salivary	Vespula Phospholipase A1 Ves m/v 1	336	AAB48072.1	897647	7
Vespula vulgaris	Wasp	Ves v 5.0101	Venom or Salivary	Vespula Ves f 5	227	AAA30333.1	162551	7
Vespula vulgaris	Wasp	Ves v 5	Venom or Salivary	Vespula Ves f 5	204	CAB42887.1	4826574	7
Vespula vulgaris	Wasp	Ves v 5	Venom or Salivary	Vespula Ves f 5	209	1QNX_A	11514279	7
Vespula vulgaris	Wasp	Ves v 2	Venom or Salivary	Vespula Ves v 2	331	P49370.1	1346323	7
Vespula vulgaris	Wasp	Ves v 2.0101	Venom or Salivary	Vespula Ves v 2	340	CAI77218.1	62147665	7
Vespula vulgaris	Wasp	Unassigned	Venom or Salivary	Vespula Ves v 2	331	2ATM_A	109157163	8
Vespula vulgaris	Wasp	Ves v 3.0101	Venom or Salivary	Vespula Ves v 3 dipeptidylpeptidase IV	776	ACA00159.1	167782086	9
Vigna radiata	mung bean	Vig r 1.0101	Food Plant	Vigna Vig r 1 PR 10	155	AAX19889.1	60418924	7
Vigna radiata	mung bean	Vig r 2.0101	Food Plant	Vigna Vig r 2	453	ABG02262.1	108743976	15
Vigna radiata	mung bean	Vig r 2.0201	Food Plant	Vigna Vig r 2	454	ABW23574.1	158251953	15
Vigna radiata	mung bean	Vig r 6.0101	Food Plant	Vigna Vig r 6 Cytokinin-specific binding protein	155	BAA74451.1	4190976	14
Vigna radiata var. radiata	mung bean	Vig r 4.0101	Food Plant	Vigna Vig r 4	272	CAA50008.1	1000708	15

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Vitis sp.	Grape	Unassigned	Food Plant	Vitis Lipid transfer protein P3	91	P80273.2	145559502	8
Vitis sp.	Grape	Vit v 1	Food Plant	Vitis Vit v 1 LTP	37	P80274.1	462719	7
Vitis sp.	Grape	Unassigned	Food Plant	Vitis Vit v 1 LTP	38	P33556.1	462717	7
Xiphias gladius	Swordfish	Xip g 1.0101	Food Animal	Xiphias Xip g 1 beta-parvalbumin	109	CAR48256.1	222352960	10
Zea mays	Corn	Unassigned	Aero Plant	Zea group 13 pollen allergen	410	ABD79096.1	89892725	7
Zea mays	Corn	Unassigned	Aero Plant	Zea group 13 pollen allergen	404	ABD79097.1	89892727	7
Zea mays	Corn	Unassigned	Aero Plant	Zea group 13 pollen allergen	411	ABD79098.1	89892729	7
Zea mays	Corn	Zea m 8.0101	Unassigned	Zea mays Zea m 8	278	ACX37090.1	260401081	17
Zea mays	Corn	Unassigned	Unassigned	Zea mays Zea m 8	280	P29022.1	116329	17
Zea mays	Corn	Unassigned	Aero Plant	Zea pollen specific protein	170	2209273A	1588669	7
Zea mays	Corn	Zea m 1.0101	Aero Plant	Zea Zea m 1 beta-expansin	269	AAO45607.1	28630919	7
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 1 beta-expansin	269	AAO45608.1	28630923	7
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 1 beta-expansin	269	AAK56124.1	14193761	8
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 1 beta-expansin	245	2HCZ_X	114794319	8
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 1 isoform	263	ABD79094.1	89892721	7
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 1 isoform	252	ABD79095.1	89892723	7
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 1 isoform	99	ABF81661.1	105969543	8
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 1 isoform	269	ABF81662.1	105969545	8
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 1 isoform	270	Q1ZYQ8.2	115502167	9
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 1 isoform	269	P0C1Y5.1	115502168	9
Zea mays	Corn	Zea m 12.0104	Aero Plant	Zea Zea m 12 profilin	131	AAB86960.1	2642324	7
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 12 profilin	131	ABG81312.1	110644952	8

Species	Common	IUIS Allergen	Type	Group*	Length	Accession	GI#@	First Version
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 12 profilin	131	ABG81313.1	110644954	8
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 12 profilin	131	ABG81314.1	110644956	8
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 12 profilin	131	ABG81315.1	110644958	8
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 12 profilin	131	ABG81316.1	110644960	8
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 12 profilin	131	ABG81317.1	110644962	8
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 12 profilin	130	ABG81318.1	110644964	8
Zea mays	Corn	Zea m 12.0101	Aero Plant	Zea Zea m 12 profilin	131	CAA51718.1	313138	15
Zea mays	Corn	Zea m 12.0102	Aero Plant	Zea Zea m 12 profilin	137	CAA51719.1	313140	15
Zea mays	Corn	Zea m 12.0103	Aero Plant	Zea Zea m 12 profilin	131	CAA51720.1	313142	15
Zea mays	Corn	Zea m 12.0105	Aero Plant	Zea Zea m 12 profilin	131	AAG35601.1	11493677	15
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 12 profilin	132	5FEF_A	1064245368	18
Zea mays	Corn	Zea m 14.0101	Food Plant	Zea Zea m 14	120	AAA33493.1	168576	15
Zea mays	Corn	Zea m 14.0102	Food Plant	Zea Zea m 14	99	AAA33494.1	168578	15
Zea mays	Corn	Unassigned	Aero Plant	Zea Zea m 25 thioredoxin	128	CAI64400.1	66841002	7
Ziziphus mauritiana	Chinese-date	Ziz m 1.0101	Food Plant	Ziziphus Ziz m 1	330	AAX40948.1	61225281	7

D.1 Omitted allergens from allergenonline

A few of the entries were omitted, due to wrong accession codes, unpublished sequences or other errors:

E Results from the EFSA scientific opinion recommended allergen analysis of *41SaM2-54* glucoamylase using allergenonline database

E.1 35% or larger identity over any 80 amino acid window

```
P53BAQ_window80_6 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_7 - 80 aa XP_003030591.1 48.750% identity in 80 aa overlap
P53BAQ_window80_8 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_10 - 80 aa XP_003030591.1 51.250% identity in 80 aa overlap
P53BAQ_window80_12 - 80 aa XP_003030591.1 52.500% identity in 80 aa overlap
P53BAQ_window80_14 - 80 aa XP_003030591.1 51.250% identity in 80 aa overlap
P53BAQ_window80_16 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_19 - 80 aa XP_003030591.1 52.500% identity in 80 aa overlap
P53BAQ_window80_23 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_24 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_25 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_26 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_27 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_28 - 80 aa XP_003030591.1 51.250% identity in 80 aa overlap
P53BAQ_window80_33 - 80 aa XP_003030591.1 55.000% identity in 80 aa overlap
P53BAQ_window80_34 - 80 aa XP_003030591.1 53.750% identity in 80 aa overlap
P53BAQ_window80_36 - 80 aa XP_003030591.1 53.750% identity in 80 aa overlap
P53BAQ_window80_37 - 80 aa XP_003030591.1 55.000% identity in 80 aa overlap
P53BAQ_window80_38 - 80 aa XP_003030591.1 55.000% identity in 80 aa overlap
P53BAQ_window80_39 - 80 aa XP_003030591.1 56.250% identity in 80 aa overlap
P53BAQ_window80_40 - 80 aa XP_003030591.1 57.500% identity in 80 aa overlap
P53BAQ_window80_42 - 80 aa XP_003030591.1 57.500% identity in 80 aa overlap
P53BAQ_window80_43 - 80 aa XP_003030591.1 57.500% identity in 80 aa overlap
P53BAQ_window80_44 - 80 aa XP_003030591.1 57.500% identity in 80 aa overlap
P53BAQ_window80_46 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_47 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_48 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_49 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_50 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_51 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_52 - 80 aa XP_003030591.1 62.500% identity in 80 aa overlap
P53BAQ_window80_53 - 80 aa XP_003030591.1 62.500% identity in 80 aa overlap
P53BAQ_window80_54 - 80 aa XP_003030591.1 63.750% identity in 80 aa overlap
P53BAQ_window80_57 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_58 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_60 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_62 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_63 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_65 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_67 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_77 - 80 aa XP_003030591.1 56.250% identity in 80 aa overlap
P53BAQ_window80_80 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_84 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_85 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_87 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_91 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_93 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_95 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_97 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_98 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_101 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
```

P53BAQ_window80_102 - 80 aa XP_003030591.1 62.500% identity in 80 aa overlap
P53BAQ_window80_103 - 80 aa XP_003030591.1 63.750% identity in 80 aa overlap
P53BAQ_window80_106 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_107 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_108 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_109 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_110 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_111 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_112 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_115 - 80 aa XP_003030591.1 63.750% identity in 80 aa overlap
P53BAQ_window80_117 - 80 aa XP_003030591.1 62.500% identity in 80 aa overlap
P53BAQ_window80_118 - 80 aa XP_003030591.1 62.500% identity in 80 aa overlap
P53BAQ_window80_121 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_125 - 80 aa XP_003030591.1 57.500% identity in 80 aa overlap
P53BAQ_window80_126 - 80 aa XP_003030591.1 57.500% identity in 80 aa overlap
P53BAQ_window80_133 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_139 - 80 aa XP_003030591.1 48.750% identity in 80 aa overlap
P53BAQ_window80_142 - 80 aa XP_003030591.1 47.500% identity in 80 aa overlap
P53BAQ_window80_143 - 80 aa XP_003030591.1 48.750% identity in 80 aa overlap
P53BAQ_window80_144 - 80 aa XP_003030591.1 48.750% identity in 80 aa overlap
P53BAQ_window80_146 - 80 aa XP_003030591.1 51.250% identity in 80 aa overlap
P53BAQ_window80_147 - 80 aa XP_003030591.1 52.500% identity in 80 aa overlap
P53BAQ_window80_150 - 80 aa XP_003030591.1 55.000% identity in 80 aa overlap
P53BAQ_window80_154 - 80 aa XP_003030591.1 54.878% identity in 82 aa overlap
P53BAQ_window80_155 - 80 aa XP_003030591.1 56.250% identity in 80 aa overlap
P53BAQ_window80_156 - 80 aa XP_003030591.1 56.250% identity in 80 aa overlap
P53BAQ_window80_157 - 80 aa XP_003030591.1 54.878% identity in 82 aa overlap
P53BAQ_window80_158 - 80 aa XP_003030591.1 54.878% identity in 82 aa overlap
P53BAQ_window80_159 - 80 aa XP_003030591.1 54.321% identity in 81 aa overlap
P53BAQ_window80_160 - 80 aa XP_003030591.1 54.878% identity in 82 aa overlap
P53BAQ_window80_161 - 80 aa XP_003030591.1 54.878% identity in 82 aa overlap
P53BAQ_window80_162 - 80 aa XP_003030591.1 56.250% identity in 80 aa overlap
P53BAQ_window80_163 - 80 aa XP_003030591.1 56.250% identity in 80 aa overlap
P53BAQ_window80_164 - 80 aa XP_003030591.1 56.098% identity in 82 aa overlap
P53BAQ_window80_165 - 80 aa XP_003030591.1 56.790% identity in 81 aa overlap
P53BAQ_window80_166 - 80 aa XP_003030591.1 57.317% identity in 82 aa overlap
P53BAQ_window80_167 - 80 aa XP_003030591.1 56.790% identity in 81 aa overlap
P53BAQ_window80_168 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_169 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_170 - 80 aa XP_003030591.1 58.537% identity in 82 aa overlap
P53BAQ_window80_171 - 80 aa XP_003030591.1 57.317% identity in 82 aa overlap
P53BAQ_window80_172 - 80 aa XP_003030591.1 56.098% identity in 82 aa overlap
P53BAQ_window80_175 - 80 aa XP_003030591.1 58.025% identity in 81 aa overlap
P53BAQ_window80_176 - 80 aa XP_003030591.1 58.537% identity in 82 aa overlap
P53BAQ_window80_177 - 80 aa XP_003030591.1 58.537% identity in 82 aa overlap
P53BAQ_window80_178 - 80 aa XP_003030591.1 58.537% identity in 82 aa overlap
P53BAQ_window80_179 - 80 aa XP_003030591.1 58.537% identity in 82 aa overlap
P53BAQ_window80_180 - 80 aa XP_003030591.1 58.537% identity in 82 aa overlap
P53BAQ_window80_181 - 80 aa XP_003030591.1 58.025% identity in 81 aa overlap
P53BAQ_window80_182 - 80 aa XP_003030591.1 57.500% identity in 80 aa overlap
P53BAQ_window80_184 - 80 aa XP_003030591.1 56.790% identity in 81 aa overlap
P53BAQ_window80_185 - 80 aa XP_003030591.1 57.317% identity in 82 aa overlap
P53BAQ_window80_186 - 80 aa XP_003030591.1 57.317% identity in 82 aa overlap
P53BAQ_window80_187 - 80 aa XP_003030591.1 56.790% identity in 81 aa overlap
P53BAQ_window80_188 - 80 aa XP_003030591.1 56.250% identity in 80 aa overlap
P53BAQ_window80_189 - 80 aa XP_003030591.1 54.878% identity in 82 aa overlap
P53BAQ_window80_190 - 80 aa XP_003030591.1 54.878% identity in 82 aa overlap
P53BAQ_window80_191 - 80 aa XP_003030591.1 54.878% identity in 82 aa overlap
P53BAQ_window80_192 - 80 aa XP_003030591.1 55.556% identity in 81 aa overlap
P53BAQ_window80_193 - 80 aa XP_003030591.1 56.098% identity in 82 aa overlap
P53BAQ_window80_194 - 80 aa XP_003030591.1 56.098% identity in 82 aa overlap
P53BAQ_window80_195 - 80 aa XP_003030591.1 56.250% identity in 80 aa overlap
P53BAQ_window80_196 - 80 aa XP_003030591.1 56.098% identity in 82 aa overlap
P53BAQ_window80_198 - 80 aa XP_003030591.1 57.500% identity in 80 aa overlap
P53BAQ_window80_199 - 80 aa XP_003030591.1 56.790% identity in 81 aa overlap
P53BAQ_window80_200 - 80 aa XP_003030591.1 57.317% identity in 82 aa overlap
P53BAQ_window80_202 - 80 aa XP_003030591.1 58.025% identity in 81 aa overlap
P53BAQ_window80_203 - 80 aa XP_003030591.1 58.537% identity in 82 aa overlap

P53BAQ_window80_204 - 80 aa XP_003030591.1 59.259% identity in 81 aa overlap
P53BAQ_window80_205 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_208 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_209 - 80 aa XP_003030591.1 61.728% identity in 81 aa overlap
P53BAQ_window80_210 - 80 aa XP_003030591.1 61.728% identity in 81 aa overlap
P53BAQ_window80_211 - 80 aa XP_003030591.1 61.728% identity in 81 aa overlap
P53BAQ_window80_212 - 80 aa XP_003030591.1 62.195% identity in 82 aa overlap
P53BAQ_window80_213 - 80 aa XP_003030591.1 62.500% identity in 80 aa overlap
P53BAQ_window80_214 - 80 aa XP_003030591.1 62.963% identity in 81 aa overlap
P53BAQ_window80_215 - 80 aa XP_003030591.1 63.415% identity in 82 aa overlap
P53BAQ_window80_216 - 80 aa XP_003030591.1 64.198% identity in 81 aa overlap
P53BAQ_window80_217 - 80 aa XP_003030591.1 65.000% identity in 80 aa overlap
P53BAQ_window80_228 - 80 aa XP_003030591.1 56.098% identity in 82 aa overlap
P53BAQ_window80_229 - 80 aa XP_003030591.1 56.098% identity in 82 aa overlap
P53BAQ_window80_231 - 80 aa XP_003030591.1 57.500% identity in 80 aa overlap
P53BAQ_window80_233 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_237 - 80 aa XP_003030591.1 58.750% identity in 80 aa overlap
P53BAQ_window80_240 - 80 aa XP_003030591.1 59.259% identity in 81 aa overlap
P53BAQ_window80_242 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_244 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_249 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_250 - 80 aa XP_003030591.1 59.259% identity in 81 aa overlap
P53BAQ_window80_251 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_252 - 80 aa XP_003030591.1 62.500% identity in 80 aa overlap
P53BAQ_window80_254 - 80 aa XP_003030591.1 63.750% identity in 80 aa overlap
P53BAQ_window80_255 - 80 aa XP_003030591.1 61.728% identity in 81 aa overlap
P53BAQ_window80_256 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_258 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_263 - 80 aa XP_003030591.1 61.250% identity in 80 aa overlap
P53BAQ_window80_264 - 80 aa XP_003030591.1 60.000% identity in 80 aa overlap
P53BAQ_window80_266 - 80 aa XP_003030591.1 59.259% identity in 81 aa overlap
P53BAQ_window80_270 - 80 aa XP_003030591.1 56.790% identity in 81 aa overlap
P53BAQ_window80_271 - 80 aa XP_003030591.1 57.500% identity in 80 aa overlap
P53BAQ_window80_273 - 80 aa XP_003030591.1 56.250% identity in 80 aa overlap
P53BAQ_window80_275 - 80 aa XP_003030591.1 55.556% identity in 81 aa overlap
P53BAQ_window80_276 - 80 aa XP_003030591.1 55.000% identity in 80 aa overlap
P53BAQ_window80_280 - 80 aa XP_003030591.1 52.500% identity in 80 aa overlap
P53BAQ_window80_281 - 80 aa XP_003030591.1 52.500% identity in 80 aa overlap
P53BAQ_window80_285 - 80 aa XP_003030591.1 51.852% identity in 81 aa overlap
P53BAQ_window80_286 - 80 aa XP_003030591.1 52.500% identity in 80 aa overlap
P53BAQ_window80_288 - 80 aa XP_003030591.1 51.250% identity in 80 aa overlap
P53BAQ_window80_290 - 80 aa XP_003030591.1 49.383% identity in 81 aa overlap
P53BAQ_window80_291 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_303 - 80 aa XP_003030591.1 47.500% identity in 80 aa overlap
P53BAQ_window80_306 - 80 aa XP_003030591.1 50.000% identity in 80 aa overlap
P53BAQ_window80_307 - 80 aa XP_003030591.1 48.750% identity in 80 aa overlap
P53BAQ_window80_308 - 80 aa XP_003030591.1 48.750% identity in 80 aa overlap
P53BAQ_window80_310 - 80 aa XP_003030591.1 47.500% identity in 80 aa overlap
P53BAQ_window80_311 - 80 aa XP_003030591.1 46.250% identity in 80 aa overlap
P53BAQ_window80_316 - 80 aa XP_003030591.1 45.000% identity in 80 aa overlap
P53BAQ_window80_317 - 80 aa XP_003030591.1 43.750% identity in 80 aa overlap
P53BAQ_window80_318 - 80 aa XP_003030591.1 43.750% identity in 80 aa overlap
P53BAQ_window80_319 - 80 aa XP_003030591.1 43.750% identity in 80 aa overlap
P53BAQ_window80_321 - 80 aa XP_003030591.1 42.500% identity in 80 aa overlap
P53BAQ_window80_323 - 80 aa XP_003030591.1 43.750% identity in 80 aa overlap
P53BAQ_window80_327 - 80 aa XP_003030591.1 42.500% identity in 80 aa overlap
P53BAQ_window80_328 - 80 aa XP_003030591.1 41.250% identity in 80 aa overlap
P53BAQ_window80_330 - 80 aa XP_003030591.1 40.000% identity in 80 aa overlap
P53BAQ_window80_332 - 80 aa XP_003030591.1 38.750% identity in 80 aa overlap
P53BAQ_window80_339 - 80 aa XP_003030591.1 40.000% identity in 80 aa overlap
P53BAQ_window80_340 - 80 aa XP_003030591.1 40.000% identity in 80 aa overlap
P53BAQ_window80_341 - 80 aa XP_003030591.1 41.250% identity in 80 aa overlap
P53BAQ_window80_342 - 80 aa XP_003030591.1 41.250% identity in 80 aa overlap
P53BAQ_window80_343 - 80 aa XP_003030591.1 41.250% identity in 80 aa overlap
P53BAQ_window80_352 - 80 aa XP_003030591.1 42.500% identity in 80 aa overlap
P53BAQ_window80_359 - 80 aa XP_003030591.1 43.750% identity in 80 aa overlap
P53BAQ_window80_369 - 80 aa XP_003030591.1 42.500% identity in 80 aa overlap
P53BAQ_window80_370 - 80 aa XP_003030591.1 43.750% identity in 80 aa overlap

```
P53BAQ_window80_372 - 80 aa XP_003030591.1 45.000% identity in 80 aa overlap
P53BAQ_window80_379 - 80 aa XP_003030591.1 45.000% identity in 80 aa overlap
P53BAQ_window80_431 - 80 aa XP_003030591.1 35.000% identity in 80 aa overlap
P53BAQ_window80_436 - 80 aa XP_003030591.1 36.250% identity in 80 aa overlap
P53BAQ_window80_450 - 80 aa XP_003030591.1 35.000% identity in 80 aa overlap
P53BAQ_window80_451 - 80 aa XP_003030591.1 35.000% identity in 80 aa overlap
P53BAQ_window80_452 - 80 aa XP_003030591.1 35.366% identity in 82 aa overlap
P53BAQ_window80_462 - 80 aa XP_003030591.1 41.250% identity in 80 aa overlap
P53BAQ_window80_463 - 80 aa XP_003030591.1 41.975% identity in 81 aa overlap
P53BAQ_window80_464 - 80 aa XP_003030591.1 42.683% identity in 82 aa overlap
P53BAQ_window80_465 - 80 aa XP_003030591.1 43.373% identity in 83 aa overlap
P53BAQ_window80_466 - 80 aa XP_003030591.1 44.048% identity in 84 aa overlap
P53BAQ_window80_467 - 80 aa XP_003030591.1 44.048% identity in 84 aa overlap
P53BAQ_window80_468 - 80 aa XP_003030591.1 44.048% identity in 84 aa overlap
P53BAQ_window80_469 - 80 aa XP_003030591.1 43.678% identity in 87 aa overlap
P53BAQ_window80_470 - 80 aa XP_003030591.1 43.902% identity in 82 aa overlap
P53BAQ_window80_471 - 80 aa XP_003030591.1 43.902% identity in 82 aa overlap
P53BAQ_window80_472 - 80 aa XP_003030591.1 43.529% identity in 85 aa overlap
P53BAQ_window80_473 - 80 aa XP_003030591.1 44.186% identity in 86 aa overlap
P53BAQ_window80_474 - 80 aa XP_003030591.1 44.828% identity in 87 aa overlap
P53BAQ_window80_475 - 80 aa XP_003030591.1 43.750% identity in 80 aa overlap
P53BAQ_window80_489 - 80 aa XP_003030591.1 47.500% identity in 80 aa overlap
P53BAQ_window80_490 - 80 aa XP_003030591.1 48.750% identity in 80 aa overlap
P53BAQ_window80_491 - 80 aa XP_003030591.1 48.750% identity in 80 aa overlap
P53BAQ_window80_492 - 80 aa XP_003030591.1 48.750% identity in 80 aa overlap
P53BAQ_window80_493 - 80 aa XP_003030591.1 48.750% identity in 80 aa overlap
P53BAQ_window80_511 - 80 aa XP_003030591.1 46.250% identity in 80 aa overlap
P53BAQ_window80_512 - 80 aa XP_003030591.1 45.000% identity in 80 aa overlap
```

(Count of significant hits described in text based on identity $> 35\%$.)

E.2 35% or larger identity over any 80 amino acid window (with scaling)

```
P53BAQ_window80_1 - 80 aa XP_003030591.1 49.367% identity in 79 aa overlap
P53BAQ_window80_2 - 80 aa XP_003030591.1 49.367% identity in 79 aa overlap
P53BAQ_window80_3 - 80 aa XP_003030591.1 50.667% identity in 75 aa overlap
P53BAQ_window80_4 - 80 aa XP_003030591.1 50.000% identity in 78 aa overlap
P53BAQ_window80_5 - 80 aa XP_003030591.1 49.367% identity in 79 aa overlap
P53BAQ_window80_9 - 80 aa XP_003030591.1 50.633% identity in 79 aa overlap
P53BAQ_window80_11 - 80 aa XP_003030591.1 51.899% identity in 79 aa overlap
P53BAQ_window80_13 - 80 aa XP_003030591.1 51.899% identity in 79 aa overlap
P53BAQ_window80_15 - 80 aa XP_003030591.1 50.633% identity in 79 aa overlap
P53BAQ_window80_17 - 80 aa XP_003030591.1 51.282% identity in 78 aa overlap
P53BAQ_window80_18 - 80 aa XP_003030591.1 51.899% identity in 79 aa overlap
P53BAQ_window80_20 - 80 aa XP_003030591.1 51.899% identity in 79 aa overlap
P53BAQ_window80_21 - 80 aa XP_003030591.1 52.564% identity in 78 aa overlap
P53BAQ_window80_22 - 80 aa XP_003030591.1 50.633% identity in 79 aa overlap
P53BAQ_window80_29 - 80 aa XP_003030591.1 53.947% identity in 76 aa overlap
P53BAQ_window80_30 - 80 aa XP_003030591.1 54.545% identity in 77 aa overlap
P53BAQ_window80_31 - 80 aa XP_003030591.1 55.128% identity in 78 aa overlap
P53BAQ_window80_32 - 80 aa XP_003030591.1 55.696% identity in 79 aa overlap
P53BAQ_window80_35 - 80 aa XP_003030591.1 53.165% identity in 79 aa overlap
P53BAQ_window80_41 - 80 aa XP_003030591.1 58.228% identity in 79 aa overlap
P53BAQ_window80_45 - 80 aa XP_003030591.1 58.228% identity in 79 aa overlap
P53BAQ_window80_55 - 80 aa XP_003030591.1 63.291% identity in 79 aa overlap
P53BAQ_window80_56 - 80 aa XP_003030591.1 62.821% identity in 78 aa overlap
P53BAQ_window80_59 - 80 aa XP_003030591.1 60.759% identity in 79 aa overlap
P53BAQ_window80_61 - 80 aa XP_003030591.1 60.256% identity in 78 aa overlap
P53BAQ_window80_64 - 80 aa XP_003030591.1 62.025% identity in 79 aa overlap
P53BAQ_window80_66 - 80 aa XP_003030591.1 60.759% identity in 79 aa overlap
P53BAQ_window80_68 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_69 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_70 - 80 aa XP_003030591.1 58.974% identity in 78 aa overlap
```

P53BAQ_window80_71 - 80 aa XP_003030591.1 58.228% identity in 79 aa overlap
P53BAQ_window80_72 - 80 aa XP_003030591.1 58.228% identity in 79 aa overlap
P53BAQ_window80_73 - 80 aa XP_003030591.1 60.000% identity in 75 aa overlap
P53BAQ_window80_74 - 80 aa XP_003030591.1 60.000% identity in 75 aa overlap
P53BAQ_window80_75 - 80 aa XP_003030591.1 60.000% identity in 75 aa overlap
P53BAQ_window80_76 - 80 aa XP_003030591.1 60.000% identity in 75 aa overlap
P53BAQ_window80_78 - 80 aa XP_003030591.1 58.974% identity in 78 aa overlap
P53BAQ_window80_79 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_81 - 80 aa XP_003030591.1 60.256% identity in 78 aa overlap
P53BAQ_window80_82 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_83 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_86 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_88 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_89 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_90 - 80 aa XP_003030591.1 58.974% identity in 78 aa overlap
P53BAQ_window80_92 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_94 - 80 aa XP_003030591.1 61.538% identity in 78 aa overlap
P53BAQ_window80_96 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_99 - 80 aa XP_003030591.1 60.256% identity in 78 aa overlap
P53BAQ_window80_100 - 80 aa XP_003030591.1 60.759% identity in 79 aa overlap
P53BAQ_window80_104 - 80 aa XP_003030591.1 63.291% identity in 79 aa overlap
P53BAQ_window80_105 - 80 aa XP_003030591.1 62.821% identity in 78 aa overlap
P53BAQ_window80_113 - 80 aa XP_003030591.1 63.636% identity in 77 aa overlap
P53BAQ_window80_114 - 80 aa XP_003030591.1 63.291% identity in 79 aa overlap
P53BAQ_window80_116 - 80 aa XP_003030591.1 63.291% identity in 79 aa overlap
P53BAQ_window80_119 - 80 aa XP_003030591.1 62.025% identity in 79 aa overlap
P53BAQ_window80_120 - 80 aa XP_003030591.1 62.338% identity in 77 aa overlap
P53BAQ_window80_122 - 80 aa XP_003030591.1 60.759% identity in 79 aa overlap
P53BAQ_window80_123 - 80 aa XP_003030591.1 60.256% identity in 78 aa overlap
P53BAQ_window80_124 - 80 aa XP_003030591.1 59.740% identity in 77 aa overlap
P53BAQ_window80_127 - 80 aa XP_003030591.1 56.962% identity in 79 aa overlap
P53BAQ_window80_128 - 80 aa XP_003030591.1 56.410% identity in 78 aa overlap
P53BAQ_window80_129 - 80 aa XP_003030591.1 55.844% identity in 77 aa overlap
P53BAQ_window80_130 - 80 aa XP_003030591.1 55.263% identity in 76 aa overlap
P53BAQ_window80_131 - 80 aa XP_003030591.1 54.667% identity in 75 aa overlap
P53BAQ_window80_132 - 80 aa XP_003030591.1 54.054% identity in 74 aa overlap
P53BAQ_window80_134 - 80 aa XP_003030591.1 50.649% identity in 77 aa overlap
P53BAQ_window80_135 - 80 aa XP_003030591.1 50.649% identity in 77 aa overlap
P53BAQ_window80_136 - 80 aa XP_003030591.1 50.649% identity in 77 aa overlap
P53BAQ_window80_137 - 80 aa XP_003030591.1 50.000% identity in 76 aa overlap
P53BAQ_window80_138 - 80 aa XP_003030591.1 48.101% identity in 79 aa overlap
P53BAQ_window80_140 - 80 aa XP_003030591.1 48.718% identity in 78 aa overlap
P53BAQ_window80_141 - 80 aa XP_003030591.1 48.718% identity in 78 aa overlap
P53BAQ_window80_145 - 80 aa XP_003030591.1 50.633% identity in 79 aa overlap
P53BAQ_window80_148 - 80 aa XP_003030591.1 53.846% identity in 78 aa overlap
P53BAQ_window80_149 - 80 aa XP_003030591.1 54.430% identity in 79 aa overlap
P53BAQ_window80_151 - 80 aa XP_003030591.1 56.579% identity in 76 aa overlap
P53BAQ_window80_152 - 80 aa XP_003030591.1 56.579% identity in 76 aa overlap
P53BAQ_window80_153 - 80 aa XP_003030591.1 54.430% identity in 79 aa overlap
P53BAQ_window80_173 - 80 aa XP_003030591.1 58.228% identity in 79 aa overlap
P53BAQ_window80_174 - 80 aa XP_003030591.1 58.228% identity in 79 aa overlap
P53BAQ_window80_183 - 80 aa XP_003030591.1 58.442% identity in 77 aa overlap
P53BAQ_window80_197 - 80 aa XP_003030591.1 56.962% identity in 79 aa overlap
P53BAQ_window80_201 - 80 aa XP_003030591.1 58.228% identity in 79 aa overlap
P53BAQ_window80_206 - 80 aa XP_003030591.1 60.256% identity in 78 aa overlap
P53BAQ_window80_207 - 80 aa XP_003030591.1 60.759% identity in 79 aa overlap
P53BAQ_window80_218 - 80 aa XP_003030591.1 64.557% identity in 79 aa overlap
P53BAQ_window80_219 - 80 aa XP_003030591.1 65.789% identity in 76 aa overlap
P53BAQ_window80_220 - 80 aa XP_003030591.1 65.789% identity in 76 aa overlap
P53BAQ_window80_221 - 80 aa XP_003030591.1 65.789% identity in 76 aa overlap
P53BAQ_window80_222 - 80 aa XP_003030591.1 65.333% identity in 75 aa overlap
P53BAQ_window80_223 - 80 aa XP_003030591.1 64.865% identity in 74 aa overlap
P53BAQ_window80_224 - 80 aa XP_003030591.1 64.384% identity in 73 aa overlap
P53BAQ_window80_225 - 80 aa XP_003030591.1 63.889% identity in 72 aa overlap
P53BAQ_window80_226 - 80 aa XP_003030591.1 63.380% identity in 71 aa overlap
P53BAQ_window80_227 - 80 aa XP_003030591.1 62.857% identity in 70 aa overlap
P53BAQ_window80_230 - 80 aa XP_003030591.1 57.692% identity in 78 aa overlap
P53BAQ_window80_232 - 80 aa XP_003030591.1 58.228% identity in 79 aa overlap

P53BAQ_window80_234 - 80 aa XP_003030591.1 60.256% identity in 78 aa overlap
P53BAQ_window80_235 - 80 aa XP_003030591.1 60.256% identity in 78 aa overlap
P53BAQ_window80_236 - 80 aa XP_003030591.1 60.256% identity in 78 aa overlap
P53BAQ_window80_238 - 80 aa XP_003030591.1 58.974% identity in 78 aa overlap
P53BAQ_window80_239 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_241 - 80 aa XP_003030591.1 60.759% identity in 79 aa overlap
P53BAQ_window80_243 - 80 aa XP_003030591.1 60.759% identity in 79 aa overlap
P53BAQ_window80_245 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_246 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_247 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_248 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_253 - 80 aa XP_003030591.1 63.291% identity in 79 aa overlap
P53BAQ_window80_257 - 80 aa XP_003030591.1 61.538% identity in 78 aa overlap
P53BAQ_window80_259 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_260 - 80 aa XP_003030591.1 61.039% identity in 77 aa overlap
P53BAQ_window80_261 - 80 aa XP_003030591.1 60.256% identity in 78 aa overlap
P53BAQ_window80_262 - 80 aa XP_003030591.1 60.759% identity in 79 aa overlap
P53BAQ_window80_265 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_267 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_268 - 80 aa XP_003030591.1 59.494% identity in 79 aa overlap
P53BAQ_window80_269 - 80 aa XP_003030591.1 58.974% identity in 78 aa overlap
P53BAQ_window80_272 - 80 aa XP_003030591.1 56.962% identity in 79 aa overlap
P53BAQ_window80_274 - 80 aa XP_003030591.1 56.962% identity in 79 aa overlap
P53BAQ_window80_277 - 80 aa XP_003030591.1 55.128% identity in 78 aa overlap
P53BAQ_window80_278 - 80 aa XP_003030591.1 54.545% identity in 77 aa overlap
P53BAQ_window80_279 - 80 aa XP_003030591.1 55.263% identity in 76 aa overlap
P53BAQ_window80_282 - 80 aa XP_003030591.1 52.564% identity in 78 aa overlap
P53BAQ_window80_283 - 80 aa XP_003030591.1 53.333% identity in 75 aa overlap
P53BAQ_window80_284 - 80 aa XP_003030591.1 51.899% identity in 79 aa overlap
P53BAQ_window80_287 - 80 aa XP_003030591.1 51.899% identity in 79 aa overlap
P53BAQ_window80_289 - 80 aa XP_003030591.1 51.282% identity in 78 aa overlap
P53BAQ_window80_292 - 80 aa XP_003030591.1 52.941% identity in 68 aa overlap
P53BAQ_window80_293 - 80 aa XP_003030591.1 51.429% identity in 70 aa overlap
P53BAQ_window80_294 - 80 aa XP_003030591.1 51.429% identity in 70 aa overlap
P53BAQ_window80_295 - 80 aa XP_003030591.1 51.429% identity in 70 aa overlap
P53BAQ_window80_296 - 80 aa XP_003030591.1 51.429% identity in 70 aa overlap
P53BAQ_window80_297 - 80 aa XP_003030591.1 51.429% identity in 70 aa overlap
P53BAQ_window80_298 - 80 aa XP_003030591.1 49.333% identity in 75 aa overlap
P53BAQ_window80_299 - 80 aa XP_003030591.1 50.000% identity in 76 aa overlap
P53BAQ_window80_300 - 80 aa XP_003030591.1 49.351% identity in 77 aa overlap
P53BAQ_window80_301 - 80 aa XP_003030591.1 48.718% identity in 78 aa overlap
P53BAQ_window80_302 - 80 aa XP_003030591.1 48.718% identity in 78 aa overlap
P53BAQ_window80_304 - 80 aa XP_003030591.1 49.351% identity in 77 aa overlap
P53BAQ_window80_305 - 80 aa XP_003030591.1 49.367% identity in 79 aa overlap
P53BAQ_window80_309 - 80 aa XP_003030591.1 48.718% identity in 78 aa overlap
P53BAQ_window80_312 - 80 aa XP_003030591.1 45.570% identity in 79 aa overlap
P53BAQ_window80_313 - 80 aa XP_003030591.1 44.872% identity in 78 aa overlap
P53BAQ_window80_314 - 80 aa XP_003030591.1 43.590% identity in 78 aa overlap
P53BAQ_window80_315 - 80 aa XP_003030591.1 44.304% identity in 79 aa overlap
P53BAQ_window80_320 - 80 aa XP_003030591.1 43.038% identity in 79 aa overlap
P53BAQ_window80_322 - 80 aa XP_003030591.1 43.038% identity in 79 aa overlap
P53BAQ_window80_324 - 80 aa XP_003030591.1 44.304% identity in 79 aa overlap
P53BAQ_window80_325 - 80 aa XP_003030591.1 44.304% identity in 79 aa overlap
P53BAQ_window80_326 - 80 aa XP_003030591.1 44.156% identity in 77 aa overlap
P53BAQ_window80_329 - 80 aa XP_003030591.1 40.506% identity in 79 aa overlap
P53BAQ_window80_331 - 80 aa XP_003030591.1 39.241% identity in 79 aa overlap
P53BAQ_window80_333 - 80 aa XP_003030591.1 37.975% identity in 79 aa overlap
P53BAQ_window80_334 - 80 aa XP_003030591.1 38.961% identity in 77 aa overlap
P53BAQ_window80_335 - 80 aa XP_003030591.1 39.744% identity in 78 aa overlap
P53BAQ_window80_336 - 80 aa XP_003030591.1 39.241% identity in 79 aa overlap
P53BAQ_window80_337 - 80 aa XP_003030591.1 39.241% identity in 79 aa overlap
P53BAQ_window80_338 - 80 aa XP_003030591.1 39.241% identity in 79 aa overlap
P53BAQ_window80_344 - 80 aa XP_003030591.1 43.038% identity in 79 aa overlap
P53BAQ_window80_345 - 80 aa XP_003030591.1 43.038% identity in 79 aa overlap
P53BAQ_window80_346 - 80 aa XP_003030591.1 43.590% identity in 78 aa overlap
P53BAQ_window80_347 - 80 aa XP_003030591.1 43.590% identity in 78 aa overlap
P53BAQ_window80_348 - 80 aa XP_003030591.1 44.304% identity in 79 aa overlap
P53BAQ_window80_349 - 80 aa XP_003030591.1 44.304% identity in 79 aa overlap

P53BAQ_window80_350 - 80 aa XP_003030591.1 43.590% identity in 78 aa overlap
P53BAQ_window80_351 - 80 aa XP_003030591.1 43.421% identity in 76 aa overlap
P53BAQ_window80_353 - 80 aa XP_003030591.1 41.772% identity in 79 aa overlap
P53BAQ_window80_354 - 80 aa XP_003030591.1 43.836% identity in 73 aa overlap
P53BAQ_window80_355 - 80 aa XP_003030591.1 43.836% identity in 73 aa overlap
P53BAQ_window80_356 - 80 aa XP_003030591.1 41.558% identity in 77 aa overlap
P53BAQ_window80_357 - 80 aa XP_003030591.1 42.308% identity in 78 aa overlap
P53BAQ_window80_358 - 80 aa XP_003030591.1 43.038% identity in 79 aa overlap
P53BAQ_window80_360 - 80 aa XP_003030591.1 45.455% identity in 77 aa overlap
P53BAQ_window80_361 - 80 aa XP_003030591.1 46.154% identity in 78 aa overlap
P53BAQ_window80_362 - 80 aa XP_003030591.1 46.154% identity in 78 aa overlap
P53BAQ_window80_363 - 80 aa XP_003030591.1 46.154% identity in 78 aa overlap
P53BAQ_window80_364 - 80 aa XP_003030591.1 45.455% identity in 77 aa overlap
P53BAQ_window80_365 - 80 aa XP_003030591.1 44.737% identity in 76 aa overlap
P53BAQ_window80_366 - 80 aa XP_003030591.1 44.595% identity in 74 aa overlap
P53BAQ_window80_367 - 80 aa XP_003030591.1 44.595% identity in 74 aa overlap
P53BAQ_window80_368 - 80 aa XP_003030591.1 44.444% identity in 72 aa overlap
P53BAQ_window80_371 - 80 aa XP_003030591.1 44.872% identity in 78 aa overlap
P53BAQ_window80_373 - 80 aa XP_003030591.1 47.436% identity in 78 aa overlap
P53BAQ_window80_374 - 80 aa XP_003030591.1 47.436% identity in 78 aa overlap
P53BAQ_window80_375 - 80 aa XP_003030591.1 47.436% identity in 78 aa overlap
P53BAQ_window80_376 - 80 aa XP_003030591.1 48.684% identity in 76 aa overlap
P53BAQ_window80_377 - 80 aa XP_003030591.1 48.684% identity in 76 aa overlap
P53BAQ_window80_378 - 80 aa XP_003030591.1 48.000% identity in 75 aa overlap
P53BAQ_window80_380 - 80 aa XP_003030591.1 45.570% identity in 79 aa overlap
P53BAQ_window80_381 - 80 aa XP_003030591.1 46.753% identity in 77 aa overlap
P53BAQ_window80_382 - 80 aa XP_003030591.1 46.753% identity in 77 aa overlap
P53BAQ_window80_383 - 80 aa XP_003030591.1 48.000% identity in 75 aa overlap
P53BAQ_window80_384 - 80 aa XP_003030591.1 48.000% identity in 75 aa overlap
P53BAQ_window80_385 - 80 aa XP_003030591.1 47.297% identity in 74 aa overlap
P53BAQ_window80_386 - 80 aa XP_003030591.1 46.575% identity in 73 aa overlap
P53BAQ_window80_387 - 80 aa XP_003030591.1 47.222% identity in 72 aa overlap
P53BAQ_window80_388 - 80 aa XP_003030591.1 47.143% identity in 70 aa overlap
P53BAQ_window80_389 - 80 aa XP_003030591.1 47.143% identity in 70 aa overlap
P53BAQ_window80_390 - 80 aa XP_003030591.1 50.000% identity in 66 aa overlap
P53BAQ_window80_391 - 80 aa XP_003030591.1 50.000% identity in 66 aa overlap
P53BAQ_window80_392 - 80 aa XP_003030591.1 50.000% identity in 66 aa overlap
P53BAQ_window80_393 - 80 aa XP_003030591.1 50.000% identity in 66 aa overlap
P53BAQ_window80_394 - 80 aa XP_003030591.1 50.769% identity in 65 aa overlap
P53BAQ_window80_395 - 80 aa XP_003030591.1 50.000% identity in 64 aa overlap
P53BAQ_window80_396 - 80 aa XP_003030591.1 49.206% identity in 63 aa overlap
P53BAQ_window80_397 - 80 aa XP_003030591.1 50.000% identity in 62 aa overlap
P53BAQ_window80_398 - 80 aa XP_003030591.1 49.180% identity in 61 aa overlap
P53BAQ_window80_399 - 80 aa XP_003030591.1 49.153% identity in 59 aa overlap
P53BAQ_window80_400 - 80 aa XP_003030591.1 49.153% identity in 59 aa overlap
P53BAQ_window80_401 - 80 aa XP_003030591.1 48.276% identity in 58 aa overlap
P53BAQ_window80_432 - 80 aa XP_003030591.1 36.364% identity in 77 aa overlap
P53BAQ_window80_454 - 80 aa XP_003030591.1 41.429% identity in 70 aa overlap
P53BAQ_window80_455 - 80 aa XP_003030591.1 41.429% identity in 70 aa overlap
P53BAQ_window80_456 - 80 aa XP_003030591.1 41.429% identity in 70 aa overlap
P53BAQ_window80_457 - 80 aa XP_003030591.1 41.429% identity in 70 aa overlap
P53BAQ_window80_458 - 80 aa XP_003030591.1 41.429% identity in 70 aa overlap
P53BAQ_window80_459 - 80 aa XP_003030591.1 40.260% identity in 77 aa overlap
P53BAQ_window80_460 - 80 aa XP_003030591.1 40.260% identity in 77 aa overlap
P53BAQ_window80_461 - 80 aa XP_003030591.1 40.506% identity in 79 aa overlap
P53BAQ_window80_476 - 80 aa XP_003030591.1 51.515% identity in 66 aa overlap
P53BAQ_window80_477 - 80 aa XP_003030591.1 50.000% identity in 68 aa overlap
P53BAQ_window80_478 - 80 aa XP_003030591.1 50.000% identity in 68 aa overlap
P53BAQ_window80_479 - 80 aa XP_003030591.1 50.000% identity in 68 aa overlap
P53BAQ_window80_480 - 80 aa XP_003030591.1 50.000% identity in 68 aa overlap
P53BAQ_window80_481 - 80 aa XP_003030591.1 50.000% identity in 68 aa overlap
P53BAQ_window80_482 - 80 aa XP_003030591.1 50.000% identity in 68 aa overlap
P53BAQ_window80_483 - 80 aa XP_003030591.1 50.000% identity in 68 aa overlap
P53BAQ_window80_484 - 80 aa XP_003030591.1 46.667% identity in 75 aa overlap
P53BAQ_window80_485 - 80 aa XP_003030591.1 46.667% identity in 75 aa overlap
P53BAQ_window80_486 - 80 aa XP_003030591.1 46.753% identity in 77 aa overlap
P53BAQ_window80_487 - 80 aa XP_003030591.1 47.436% identity in 78 aa overlap
P53BAQ_window80_488 - 80 aa XP_003030591.1 47.436% identity in 78 aa overlap

```
P53BAQ_window80_494 - 80 aa XP_003030591.1 51.351% identity in 74 aa overlap
P53BAQ_window80_495 - 80 aa XP_003030591.1 51.316% identity in 76 aa overlap
P53BAQ_window80_496 - 80 aa XP_003030591.1 51.316% identity in 76 aa overlap
P53BAQ_window80_497 - 80 aa XP_003030591.1 51.316% identity in 76 aa overlap
P53BAQ_window80_498 - 80 aa XP_003030591.1 50.633% identity in 79 aa overlap
P53BAQ_window80_499 - 80 aa XP_003030591.1 50.633% identity in 79 aa overlap
P53BAQ_window80_500 - 80 aa XP_003030591.1 51.282% identity in 78 aa overlap
P53BAQ_window80_501 - 80 aa XP_003030591.1 51.316% identity in 76 aa overlap
P53BAQ_window80_502 - 80 aa XP_003030591.1 51.316% identity in 76 aa overlap
P53BAQ_window80_503 - 80 aa XP_003030591.1 50.667% identity in 75 aa overlap
P53BAQ_window80_504 - 80 aa XP_003030591.1 50.000% identity in 74 aa overlap
P53BAQ_window80_505 - 80 aa XP_003030591.1 50.000% identity in 72 aa overlap
P53BAQ_window80_506 - 80 aa XP_003030591.1 50.000% identity in 72 aa overlap
P53BAQ_window80_507 - 80 aa XP_003030591.1 49.296% identity in 71 aa overlap
P53BAQ_window80_508 - 80 aa XP_003030591.1 48.571% identity in 70 aa overlap
P53BAQ_window80_509 - 80 aa XP_003030591.1 50.000% identity in 68 aa overlap
P53BAQ_window80_510 - 80 aa XP_003030591.1 50.000% identity in 68 aa overlap
```

(blank=No matches found) Count of significant hits described in text based on identity > 35%.

Esben Friis
Luna# 2018-06838-01 <epf@novozymes.com>

May 17, 2018

Allergen/Toxin Risk Assessment

By MITH, EPF
18-May-2018

Glucoamylase produced by 41SaM2-54

Sequence homology assessment:

Allergen risk assessment analysis of Glucoamylase produced by 41SaM2-54 was performed according to the EFSA scientific opinion using allergen on-line and allergen.org databases. The analyses of Glucoamylase produced by 41SaM2-54's sequence identified 1 fungal allergen having an identity with a known allergen above the threshold of 35% across a 80 amino acid window.

The glucoamylase Sch c 1 of *Schizophyllum commune* was revealed to have up to 47.9% identity with Glucoamylase produced by 41SaM2-54' 80 amino acids windows using the allergen.org database.

Similarities with other allergens were below 35%. Since it is generally accepted that proteins with such a low identity rarely share epitopes they were not considered in this report (Pearson, 2000; Aalberse et al., 2001; Hileman et al., 2002; Ladics et al., 2007).

Ladics and colleagues compared the 80 aa sliding window search method to the conventional homology sequence comparison, the latter being the overall sequence homology between two proteins over the full stretch of the protein. Results showed that the conventional FASTA alignment was a more reliable search approach when testing sequence homology to known allergens using 35% as the threshold in both methods. Conventional FASTA yields a much lower number of false positive, but the same number of false negatives (Ladics et al., 2007). Because of the high number of false positives in the 80 aa sliding window search approach with a sequence homology threshold of 35%, more and more experts suggest that this threshold may be too conservative and will overpredict cross-reactivity, which may result in the blocking of potentially useful products. There seem to be consensus that the risk of cross reactivity is considered to be very low beneath 50% homology, and low to moderate between 50% and 70% and high above 70% homology (Goodman et al., 2008, Hileman et al., 2002, Stadler et al., 2003 and Aalberse et al., 2000). In 2008, Goodman and colleagues suggested to raise the threshold of the 80 aa sliding window search to 50% sequence homology.

Allergen risk assessment:

1. Respiratory sensitization and asthma

Like fungal α-amylases (prevalence: 23%), fungal-derived glucoamylases also have been reported to trigger respiratory sensitization and asthma in bakers (prevalence: 8%) (Sander et al., 1998; Quirce et al., 2002). Further, in an occupational setting, the sensitizing potential of enzymes after respiratory

exposure, and the relation to enzyme specific asthma, has been recognized for many years (reviewed in Sarlo., 2003).

These observations mandate that Glucoamylase produced by 41SaM2-54 must be handled according to established safety guidelines for enzymes.

2. Food allergy

Sch c 1 is described as food allergen (http://www.allergen.org/).

There is compelling evidence that majority of adults affected by food-induced occupational asthma can ingest the allergen without symptoms, suggesting that inhalation is not likely to result in food allergy (Armentia et al., 2009).

In addition, the Association of Manufacturers and Formulators of Enzyme Products (AMFEP) Working Group on Consumer Allergy Risk from Enzyme Residues in Food performed an in-depth analysis of the allergenicity of enzyme products (Dauvrin et al, 1998). In this paper, Dauvrin et al (1998) conclude that enzyme exposure by ingestion, in opposition to exposure by inhalation, is extremely unlikely to lead to sensitization. Only one single case has been reported in the literature and this case was not verified as a *bona fide* oral sensitization to enzymes in food.

This is backed up by a study using the generally recognized guidelines for food allergy diagnosis (skin prick test, specific serum IgE and DBPCFC). This study included 400 patients with diagnosed allergy to one or more of inhalation allergens, food allergens, bee or wasp allergens. The study concluded that no cases of IgE-mediated food allergy to commercial enzymes (incl. proteases) could be found. There were further no indications of cross-reactivity between the tested enzymes used in food and the main known allergens represented by the patients included in the study (Bindslev-Jensen et al., 2006).

Toxin risk assessment were performed as follows. Protein sequences that contain the word toxin in the description field were extracted from UNIPROT (Database date: 2018-04-25). ClustalW 2.0.10 were used to align these sequences to the sequence of Glucoamylase produced by 41SaM2-54 and the identity percentages were calculated to the Glucoamylase produced by 41SaM2-54 sequence or the compared toxin sequence is calculated, whichever is longest. No toxins were found with high homology to Glucoamylase produced by 41SaM2-54.

See report "Glucoamylase produced by 41SaM2-54. Assessment of sequence homology to known toxins and allergens"/[LUNA # 2018-06836] for details.

Conclusion:

On the basis of the available evidence it can be concluded that oral intake of Glucoamylase produced by 41SaM2-54 is not anticipated to pose any food allergenic concern.

However, Glucoamylase produced by 41SaM2-54 should, like any enzyme, be handled according to the standard safety guidelines for enzymes both internally at Novozymes and in customer settings to avoid respiratory sensitization.

References:

- Aalberse, R.C. Structural biology of allergens, *J. Allergy Clin. Immunol.* 2000, *106*, 228-238.
- Aalberse, R., C., Akkerdaas, J., van Ree, R. 2001. Allergy 56; 478-490.
- Armentia, A., Díaz_Perales, A., Castrodeza, J., Dueñas-Laita, A., Palacin, A., Fernández S. 2009. Allergol. Immunopathol. 37; 203-204
- Bindslev-Jensen, C., Skov, P., S., Roggen, E., L., Hvass, P., Brinch, D., S. 2006. Food and Chem. Toxicol. 44; 1909-1915.
- Dauvrin T, Groot G, Maurer K-H, de Rijke D, Ryssov-Nielson H, Simonsen M, Sorensen TB (1998) Working group on consumer allergy risk from enzyme residues in food. Amfep.
- Hileman, R., E., Silvanovitch, A., Goodman, R., E., Rice, E., A., Holleschak, G., Ashwood, J.,D., Hefle, S., L. 2002. Int. Arch. Allergy Immunol. 128; 280-291.
- Ladics, G., S., Bannon, G., A., Silvanovitch, A., Cressman, R., F. 2007. Mol. Nutr. Food Res. 51; 985-998.
- Pearson, W., R. 2000. Methods Mol. Biol. 132; 185-219.
- Quirce, S., Fernández-Nieto, M., Bartolomé, B., Bombín, C., Cuevas, M., Sastre, J. 2002. Ann. Allergy Asthma Immunol. 89; 197-202.
- Sander, I., Raulf-Heimsoth, M., Siethoff, C., Lohaus, C., Meyer, H., E., Baur, X. 1998. J. Allergy Clin. Immunol. 102; 256-264.
- Sarlo K., Ann Allergy Asthma Immunol. 2003 May;90(5 Suppl 2):32-4.
- Stadler, M.B., Stadler, B.M. Allergenicity prediction by protein sequence, *FASEB J* 2003, *17*, 1141-1143.

Toxicology

Date : 09 July 2013
File : 2013-10425-01
Ref.: TrGQ/PBjP

SUMMARY OF TOXICITY DATA

Amyloglucosidase

Authors:
Trine Østergaard Lund
Peder Bjarne Pedersen

Issued by:
Novozymes A/S
Krogshoejvej 36
DK-2880 Bagsvaerd
Denmark

C O N T E N T S

PAGE

1. ABSTRACT

The below series of toxicological studies were undertaken to evaluate the safety of Amyloglucosidase, represented by batch PPY 24900 (named T-AMG, PPY 24900 in the study report) and PPY 32789 (named Amyloglucosidase or Amyloglycosidase PPY 32789 in the study reports).

All studies were carried out in accordance with current OECD guidelines and in compliance with the OECD principles of Good Laboratory Practice (GLP). The studies were performed at Novozymes A/S, Denmark, Covance Laboratories, England, Jai Research Foundation, India and at Huntingdon Life Sciences, England during the period November 2005 to June 2012.

The main conclusions of the safety studies can be summarized as follows:

- Amyloglucosidase PPY 32789 did not induce gene mutations in the Ames test, neither in the presence or absence of S-9 mix.
- Amyloglucosidase PPY 32789 did not show any clastogenic activity, neither in the presence or absence of S-9 mix, when tested in the *in vitro* micronucleus assay.
- Amyloglucosidase PPY 32789 up to the dose level of 10.0 mL/kg bw (corresponding to 4084.6 AGU/kg bw/day or 1466 mg TOS/kg bw/day) did not produce any adverse effect when administered to Wistar rats through oral gavage for 14 consecutive days.
- Daily oral administration (by gavage) of T-AMG, PPY 24900 to rats at dosages of up to 10.0 mL/kg bw/day for thirteen weeks resulted in no treatment-related effects. Consequently, the No Observed Adverse Effect level (NOAEL) was considered to be 10 mL/kg/day (equivalent to 5528 AGU/kg bw/day or 1470 mg TOS/kg/ day).

Based on the present toxicity data and the fact that Amyloglucosidases have a history of safe use, it is our conclusion that Amyloglucosidase, represented by batch PPY 24900 and PPY 32789, can be considered as generally safe.

2. TEST SUBSTANCE

Amyloglucosidase is a liquid enzyme concentrate containing an Glucan 1,4 alpha-glucosidase (AMG) (E.C. number 3.2.1.3) which hydrolysis 1,4-alpha and 1,6-alpha linkages of dextrins from non-reducing end.

2.1 Production organism

Amyloglucosidase is produced by a strain of *Aspergillus niger*, containing an Amyloglucosidase gene code which is a originating from *Rasamsonia emersoni.*.

Aspergillus niger has a long history of safe use. This species has been used for decades in the production of enzymes, and for more than a decade as recombinant organism for production of a variety of bio-industrial products.

Strains of *Aspergillus niger* may have the potential to produce certain mycotoxins of medical importance. The production strain of Amyloglucosidase belongs to a strain

lineage which has been thoroughly investigated for its potential to produce secondary metabolites. The results showed, that this strain lineage does not produce any known mycotoxins.

The test substance does not contain the production strain. Absence of the production strain is part of the complete specification of the product.

2.2 Characterization

Two batches, PPY 24900 and PPY 32789, were used for the conduct of the toxicological studies. The characterization of the batches are presented in Table 1.

Table 1.
Characterization data of Glucose isomerase, batch PPV 31321 and batch PPG 1711

Batch number	PPY 24900	PPY 32789
Activity	524 AGU/g	392 AGU/g
Water KF (% w/w)	85.6	89.8
Dry matter (% w/w)	14.4	10.2
Ash (% w/w)	0.5	0.9
Total Organic Solids (TOS[1]) (% w/w)	13.9	9.3
Specific gravity (g/mL)	1.055	1.042

[1] % TOS is calculated as 100% - % water - % ash - % diluents.

3. MUTAGENICITY

3.1 Bacterial Reverse Mutation assay (Ames test)

Amyloglucosidase PPY 32789 was examined for mutagenic activity in the bacterial reverse mutation assay using *Salmonella typhimurium* strain TA1535, TA100, TA1537, and TA98 and *Escherichia coli* WP2uvrA. The study was carried out according to the OECD test guideline 471 (adopted in 1997) and in compliance with GLP.

Crude enzyme preparations, like the present batch of Amyloglucosidase, contain the free amino acids histidine and tryptophan, most often in an amount, which exceeds the critical concentration for incorporation in the direct standard assay.
To overcome this problem all strains were exposed to the test substance in liquid culture (“treat and plate assay”).

Two independent experiments were performed, with and without the inclusion of metabolic activation (S-9 mix). In each experiment, cultures of bacteria were exposed to six doses of the test substance (5000, 2500, 1250, 625, 313, and 156 µg dry matter/mL) in a phosphate buffered nutrient broth for 3 hours. After incubation, the test substance was removed by centrifugation prior to plating.

No treatments of any of the bacterial strains with the test substance resulted in any increases in number of revertants meeting the criteria for a positive or equivocal result, either in the presence or absence of S-9 mix.

The results obtained with the diagnostic mutagens and the solvent control demonstrated the sensitivity of the tests and the efficacy of the S-9 mix metabolic activation system.

It was concluded that Amyloglucosidase PPY 32789 did not induce gene mutations in bacteria either in the presence or absence of metabolic activation when tested under the conditions employed in this study.

3.2 *In vitro* micronucleus assay

In order to assess the clastogenic and the aneugenic activity of Amyloglucosidase PPY 32789, its effects on the frequency of micronuclei was investigated in cultured human peripheral blood lymphocytes applying the cytokinesis-block methodology.

The study was conducted according to GLP, in compliance with the OECD test guideline 487 (adopted in 2010).

Heparinized whole blood cultures, pooled from two female donors, were established, and division of the lymphocytes was stimulated by adding phytohaemagglutinin (PHA) to the cultures.

Two independent experiments were performed. Sets of duplicate cultures were treated with the solvent (purified water), test substance or appropriate positive controls. Treatments with the test substance covered a broad range of doses, separated by narrow intervals. The highest concentrations used was 5000 µg/mL (expressed in terms of the test substance as supplied), which is the highest dose level recommended in the guidelines for *in vitro* cytogenetic assays.

Sets of duplicate cultures were exposed to the test substance for 3 hours in the presence and absence of metabolic activation (S-9 mix) and harvested 24 hours after the beginning of treatment (3+21 hour treatment). Additionally, a continuous 24-hour treatment without S-9 mix was included with harvesting 24 hours after removal og the test substance (24+24 hour treatment). The cultures were treated with cytochalasin-B after removal of the test substance. Three concentrations, covering an appropriate range of cytotoxicity, were selected for scoring of micronuclei by evaluating the effect of the test substance on the replication Index (RI). 2000 cells per concentration (1000 cells from each replicate culture) were scored.

The proportion of binucleate cells with micronuclei in all cultures of the vehicle controls (purified water) was within the limits of the historical ranges. The positive controls induced statistically significant increases in the proportion of cells with micronuclei, thus demonstrating the sensitivity of the test procedure and the metabolic activity of the S-9 mix employed.

Treatment of the cells with the test substance resulted in frequencies of micronucleated binuclear cells (MNBN cells), which were similar to and not significantly ($p \leq 0.05$) higher than those observed in concurrent vehicle controls for the majority of concentrations analysed. The only exeptions to this were small isolated, but statistically significant, increases observed at the lowest concentration (3000 µg/mL) following 3+21 hour treatment in the presence of S-9 and at the intermediate dose level (4000 µg/mL) after 24+24 hour treatment in the absence of S-9. However, as the MNBN cell frequencies fell within historical vehicle control (normal) values and as the higher and lower concentrations analysed showed mean MNBN cell values similar to those observed for the vehicle control, these increases were not considered to be of biological importance.

It was concluded that Amyloglucosidase PPY 32789 did not induce micronuclei in cultured human peripheral blood lymphocytes either in the absence or presence of S-9 mix under the experimental conditions employed for this study.

4. GENERAL TOXICITY

4.1 Repeated dose 14-day toxicity study in rats

In order to determine the potential systemic toxic effects of Amyloglucosidase PPY 32789 in Wistar rats, when administered through oral gavage for a period of 14 consecutive days, a repeated dose 14-day toxicity study in rats was performed.

The study was performed in compliance with GLP and the method followed was inspired from the OECD guideline 407 (2008).

Four groups each of five male and five female Wistar rats were dosed orally by gavage with Amyloglucosidase PPY 32789 at dosages of 1.0, 3.3 or 10.0 mL/kg bw (corresponding to 408.46, 1347.9 and 4084.6 AGU/kg bw/day or 147, 484 and 1466 mg TOS/kg bw/day). A similarly constituted group received the vhicle (RO water) and served as negative control. A fix dose volume of 10 mL/kg bw was used.

No mortality or treatment related clinical signs were observed during the study period.

No significant changes were observed in mean body weight, percent body weight change and mean feed consumption of male and female rats treated with the test substance compared with the control group.

No treatment related significant changes were observed in haematology and clinical chemistry of test substance treated groups compared with the control group.

No treatment related changes were observed in the terminal body weights, absolute organ weights and relative organ weights of animals from any of the treated groups compared to the control group.

External and visceral gross pathological examination of all male and female animals belonging to the control and various treatment groups did not reveal any lesion of pathological significance.

It was concluded that Amyloglucosidase PPY 32789 up to the dose level of 10.0 mL/kg bw (corresponding to 4084.6 AGU/kg bw/day or 1466 mg TOS/kg bw/day) did not produce any adverse effect when administered through oral gavage for 14 consecutive days in Wistar rats under the conditions and procedures followed in the present study.

4.2 13 week Dietary Toxicity Study in Rats

The study was carried out in accordance with the OECD guideline 408 (adopted on September 1998). It was conducted in accordance GLP.

Three groups each of ten male and ten female CH:CD® (SD) BR rats received T-AMG, batch PPY 24900, orally by gavage at dosages of 1.0, 3.3 or 10.0 mL/kg bw (equivalent to 553, 1824 and 5528 AGU/kg bw/day or 150, 480 and 1470 mg TOS/kg bw/day) for thirteen weeks. A similarly constituted group received the vehicle (purified water) and served as the negative control. The dose volume was 10 mL/kg bw for all animals.

Clinical signs were recorded daily. Detailed clinical observations were performed once weekly. During week 12 of the study, the animals were examined for sensory reactivity, grip strength and motor activity. Body weights and food consumption were recorded once weekly. Water consumption was recorded throughout the treatment period.

Ophthalmoscopy was performed on all animals before start of treatment, and on the animals of groups 1 (controls) and 4 (high dose group) during week 13 of the study. Before termination of treatment, blood samples were taken for haematology and clinical chemistry, and urine was collected for urinalysis. The animals were euthanized and subjected to a macroscopic necropsy. Specified organs and tissues were weighed, fixed and prepared for microscopic examination.

Analysis of achieved concentration of dose formulations was performed on samples taken once during weeks 1, 6 and 13. Achieved concentration was evaluated by measurement of enzyme activity.

The oral administration of T-AMG, batch PPY 24900, for 13 weeks was well- tolerated with no treatment related deaths or overt signs of toxicity. One male in the intermediate group was killed for welfare reasons during week 3. The cause of death was identified as meningitis and this was judged as not related to the treatment.

No treatment-related adverse signs were recorded at the clinical and behavioral examinations.

Bodyweight gain, food and water consumption and food conversion efficacy were unaffected by treatment.

There were no treatment-related changes on parameters for haematology, serum biochemistry, ophthalmoscopic examinations or on organ weights.

Necropsy and the following microscopic examination revealed no treatment-related effects apart from an increased incidence or increased severity of cortical vacuolation in the adrenal glands of males given 3.3 or 10.0 mL/kg bw/day. The incidence of cortical vacuolation in the present study was, however, within the ranges observed in the control CD rats from other 13-week toxicity studies at the present contract research organisation. As there were no evidence for other degenerative changes, such as inflammation or necrosis, and females were unaffected, this was considered a slight exacerbation of a normal background finding in young CD rats which was not considered toxicologically significant.

The results from the content check analysis showed that in the samples from Week 1, 6 and 13, there were no significant differences between the intended activity and the results.

In conclusion daily oral administration (by gavage) of T-AMG to rats at dosages of up to 10.0 mL/kg bw/day for thirteen weeks resulted in no treatment-related effects. Consequently, the No Observed Adverse Effect level (NOAEL) was considered to be 10 mL/kg/day (equivalent to 5528 AGU/kg bw/day or 1470 mg TOS/kg/ day).

5. CONCLUSION

Based on the present toxicity data and the fact that Amyloglucosidases have a history of safe use, it is our conclusion that Amyloglucosidase, represented by batch PPY 24900 and PPY 32789, can be considered as generally safe.

6. REFERENCES

6.1 Study reports

Novozymes A/S: Study No.: 20118069. Amyloglucosidase PPY 32789: Test for mutagenic activity with strains of *Salmonella typhimurium* and *Escherichia coli.* (December 2011). LUNA file: 2011-30186.

Covance Laboratories: Study No.: 8259272. Novozymes Reference No.: 20126003: Amyloglucosidase PPY 32789: Induction of micronuclei in cultured human peripheral blood lymphocytes. (June 2012). LUNA file: 2012-10374.

Jai Research Foundation: Study No. 410-1-02-3552. Novozymes Reference No.: 20116041: Repeated dose 14-day toxicity study of Amyloglucosidase through oral gavage in Wistar rats. (January 2012). LUNA file: 2012-03733.

Huntingdon Life Sciences: Study No.: NVZ20028/053878. Novozymes Reference No.: 20056035: T-AMG, PPY 24900 Toxicity study by Oral administration to CD Rats for 13 Weeks. (October 2006). LUNA file: 2006-46592.

6.2 Guidelines

OECD, Guidelines for testing of Chemicals. Section 3 and 4: Health effects. Organisation for Economic Co-operation and Development, Paris.

OECD principles of Good Laboratory Practice (GLP) (as revised in 1997), ENV/MC/CHEM(98)17. OECD, Paris.

Toxicology

Date: 20 December 2011
Project: OPT 05330
File: 2011-30186-01
Ref.: UF/PBjP

REPORT

Amyloglucosidase PPY32789: Test for Mutagenic Activity with Strains of *Salmonella typhimurium* and *Escherichia coli.*

Study No. 20118069

Author :
P. B. Pedersen

Issued by :
Novozymes A/S
Krogshøjvej 36
DK- 2880 Bagsværd
Denmark

Table of Contents

GLP - Compliance Statement

REPORT: Amyloglucosidase, PPY32789: Test for Mutagenic Activity with Strains of *Salmonella typhimurium* and *Escherichia coli* .

STUDY No.: 20118069

A sample of Amyloglucosidase, Batch Number: PPY32789 was received from Recovery Pilot Plant, Novozymes A/S.

This study was conducted at the department of Toxicology, Novozymes A/S in compliance with the following current Good Laboratory Practice Regulations:

OECD, ENV/MC/CHEM(98)17, 1998

Date: 21 Dec. 2011

Peder Bjarne Pedersen
Study Director

Quality assurance statement

REPORT:	Amyloglucosidase PPY32789 Test for Mutagenic Activity with Strains of Samonella typhimurium and Escherichia coli.
STUDY NUMBER	20118069

The conduct of this study has been subject to appropriate inspections and the report has been reviewed according to the relevant Standard Operation Procedures of Novozymes A/S Quality Assurance.

Inspection/Audit	Dates of inspection	Inspection results reported to Study Director and Study Management
Protocol	26 OCT 2011	26 OCT 2011
Genotype control	8 NOV 2011	8 NOV 2011
Report	14 DEC 2011	14 DEC 2011

I hereby confirm that the report reflects the raw data.

21 Dec. 2011
Date

Annie Christensen
Quality Assurance

1. General Information

STUDY

Amyloglucosidase, PPY32789: Test for Mutagenic Activity with Strains of *Salmonella typhimurium* and *Escherichia coli*.
Study No. 20118069

STUDY DIRECTOR

Peder Bjarne Pedersen
Toxicology
Novozymes A/S
Krogshøjvej 36
DK - 2880 Bagsværd

MANAGEMENT/ SPONSOR

Ditte Sidelmann Brinch
Toxicology
Novozymes A/S
Krogshøjvej 36
DK - 2880 Bagsvaerd

TEST FACILITIES

Toxicology
Novozymes A/S
Krogshøjvej 36
DK - 2880 Bagsvaerd

ARCHIVE

QM Central Archive
Novozymes A/S
Krogshøjvej 36
DK - 2880 Bagsvaerd

DATES OF STUDY

Study initiation date:	26 October 2011
Experimental start date:	27.October 2011
Experimental termination:	17.November 2011

PERSONNEL INVOLVED IN THE STUDY

Marianne S. Bønnerup (MnBq) – Toxicology
Jytte Nordlund (JNP) – Toxicology
Pia Schock (PScK) - Toxicology
Ulla Festersen (UF) - Toxicology

DATE OF FINAL REPORT

Date: 21 Dec. 2011

Peder B. Pedersen
Toxicology

2. Summary

Amyloglucosidase (Batch Number: PPY32789) was examined for mutagenic activity in the bacterial reverse mutation assay using *Salmonella typhimurium* strain TA1535, TA100, TA1537, TA98 and *Escherichia coli* WP2uvrApKM101.

Crude enzyme preparations, like the present batch of Amyloglucosidase contain the free amino acid histidine and tryptophan, most often in an amount, which exceeds the critical concentration for incorporation in the direct standard assay.
To overcome this problem all strains were exposed to Amyloglucosidase in liquid culture ("treat and plate assay"). Bacteria were exposed to 6 doses of the test substance in a phosphate buffered nutrient broth for 3 hours with 5 mg (dry matter) per ml as highest concentration. After incubation the test substance was removed by centrifugation prior to plating.

The study was conducted with and without the metabolic activation system S9 - a liver preparation from male rats, pre-treated with Aroclor 1254, and the co-factors required for mixed function oxidase activity (S9 mix). All results were confirmed by conducting two complete and independent experiments.

Amyloglucosidase contains an abundance of various nutrients, and composes a rich growth medium to the test bacteria. These circumstances are reflected in the present study. No toxicity of the test substance to the bacteria is observed. On the contrary growth stimulation is evident in several test series as demonstrated by increases in the viable count of exposed cultures compared to the solvent control. Concomitantly weak numerical increases in the number of spontaneous revertant colonies are present in some test series. These increases are obviously without biological importance.

No treatments of any of the *Salmonella* and *E.coli* strains with Amyloglucosidase resulted in any increases in revertant numbers that meets these criteria for a positive response.

It was concluded, that the results of the experiments, described in this report, give no indication of mutagenic activity of Amyloglucosidase (Batch Number: PPY32789) in the presence or absence of metabolic activation, when tested under the conditions employed in this study.

3. Introduction

Bacterial reverse mutation assays have been recognized and used for more than three decades as a rapid, sensitive and reliable method of evaluating the mutagenic potential of chemicals. Bacterial systems offer several advantages to other test systems. They can be grown in large numbers in a short time, enabling the detection of very rare mutational events. Further, extensive knowledge of bacterial genetics has allowed the construction of special strains, which are more sensitive than wild-type strains are to a variety of agents. The reversion of bacteria from growth-dependence on a particular amino acid to growth in the absence of that amino acid is the most widely used marker in reverse-mutation assays. The genetic target is small, specific and selective, and the phenotypic effect of the reverse mutation is easily detected.
A wide range of strains within the species *Salmonella typhimurium* (Ames strains) and *Escherichia coli* have been constructed in order to make the test system more sensitive and selective to different classes of chemical mutagens.
By incorporation of the post-mitochondrial supernatant (S9) from the livers of rats pretreated with an enzyme inducer Aroclor 1254, the metabolising systems present in mammalian cells are mimicked to facilitate the detection of a wide range of promutagens.

This report describes experiments performed to assess the activity of Amyloglucosidase (Batch Number: PPY32789) in amino acid dependent strains of *Salmonella typhimurium* and *Escherichia coli* capable of detecting both induced frame-shift (TA1537 and TA98) and base-pair substitution mutations (TA1535, TA100, and E. coli WP2uvrA pKM101).
Amyloglucosidase (Batch No PPY32789) is a microbial enzyme preparation derived from submerged pure culture fermentation of a non-pathogenic and non-toxigenic strain. It contains a variety of unspent medium residues, including low concentrations of free amino acids like histidine and tryptophan.
This complexity poses several problems during mutagenicity testing in vitro. In the Ames test it composes a rich growth medium to the test bacteria, resulting in completely different and poorly defined environments of exposed cultures compared to control cultures. The main problem is the content of utilizable histidine and tryptophan in the test material, since the principle of the Ames test is the histidine auxotrophy of the *Salmonella* strains and tryptophan auxotrophy of the *E.coli* strains.
As a result, the density of the bacterial background lawn increase with increasing doses ("feeding effect") followed by dose related increases in the number of spontaneous revertant. These increases are obviously artificial.
To overcome this problem all strains applied in the present study were treated with Amyloglucosidase in liquid culture ("treat and plate assay").

The study was conducted in accordance with the general recommendations in OECD Guideline for testing of chemicals, No. 471: Bacterial Reverse Mutation Assay" (July 1997 concerning the general specifications of the test. However the exposure of test bacteria in liquid culture ("treat and plate") is not specifically described in any guidelines.

4. Materials

4.1 Test substance

Amyloglucosidase (Batch Number: PPY32789, TKS number 223/11 65) was received from Recovery Pilot Plant 07.September 2011, and immediately stored in a freezer. The substance was a brown liquid with a declared content of 10.2 % (w/w) dry matter.

A standard solution of 5% (w/v) dry matter was prepared in deionised water and sterilised by filtration. Samples were sterilised by filtration and the sterility was confirmed by plate counting. Solution was stored refrigerated and used as test substance.

4.2 Positive control substances

Chemical	Source	Lot.No.
2-Nitrofluorene (2-NF)	Aldrich-Chemie	S 43858
Acridine mutagen (ICR-191)	Sigma	049K1361
1-Methyl-3-Nitro-N-NitrosoGuanidine (MNNG)	Aldrich-Chemie	15427 LO
2-Aminoanthracene (2-AA)	Aldrich-Chemie	STBB1901

All positive control substances were dissolved in dimethyl sulphoxide (spectrophotometric grade).

4.3 Liver homogenate – S9

A commercial preparation of S9 from Aroclor 1254 induced Spraque Dawley rats was obtained from Cappel/MP Biomedicals, LLC, 29525 Fountain Parkway, Solon, Ohio 44139. Specifications of the preparation, the enzymatic properties and metabolic activation from the supplier are archived as raw data.
The tubes with S9 were received frozen in dry ice and were immediately stored in a ÷80°C ultra low freezer at Toxicology, Novozymes.

4.4 Plates

As selective substrate for reverted bacteria Vogel-Bonner medium E agar plates with 2% glucose were prepared in-house as described in Appendix 3.

All plates were stored refrigerated in closed plastic bags and examined for contamina tion and dryness before use.

4.5 Bacteria

Salmonella typhimurium

Four strains of *Salmonella typhimurium* were used:
S. typhimurium TA1535
S. typhimurium TA100
S. typhimurium TA1537
S. typhimurium TA98

All these strains contain mutations in the histidine operon, thereby imposing a requirement for histidine in the growth medium. They all contain GC base-pairs at the site of the histidine mutation, and are therefore selective for agents which react predominantly with these bases. Three mutations in the histidine operon are involved: his G 46 (TA1535 and TA100) is a missense mutation which is reverted to prototrophy by a variety of mutagens that cause base-pair substitutions.
his C 3076 (TA1537) contains a frame-shift which appears to have added a GC base-pair. This mutation is reverted for example by ICR-191 and epoxides of polycyclic hydrocarbon. his D 3052 (TA98) also contains a frame-shift mutation with a sequence of repeated GC, which is reverted with the deletion of 2 of these base-pairs. It is readily reverted by aromatic amines and derivatives.

All four strains contain the deep rough (rfa) mutation, which deletes the polysaccharide side chain of the polysaccharide coat of the bacterial cell surface. This deletion increases cell permeability to more hydrophobic substances and, furthermore, greatly decreases the pathogenicity of these organisms.

The uvrB deletion renders the strains incapable of excision repair, making them more sensitive both to the mutagenic and lethal effects of a wide variety of mutagens (e.g. poly-aromatic hydrocarbons), since the strains cannot excise DNA adducts.
These two deletions include the nitrate reductase (chl) and biotin (bio) genes also.

Strain TA98 and TA100 are derived from strain TA1538 and TA1535 respectively by the addition of a plasmid, pKM101, which confers resistance to ampicillin. This plasmid also carries a gene (muc^+), which in some strains ($recA^+/lexA^+$) have proven to participate in "SOS" DNA-repair. This repair pathway is induced by DNA damage and confers resistance to the lethal effects of many mutagens at the expense of increased mutability. Bacteria carrying pKM101 have therefore a higher spontaneous mutation rate.

Escherichia coli

One strain was used:
Escherichia coli WP2uvrApKM101

This strain contain an ochre mutation in the trpE locus and can be mutated to tryptophan independence either by a base-pair reversion of an A-T base-pair in the trpE locus, or more likely, by a base-pair substitution within a number of transfer RNA loci elsewhere in the chromosome. The latter causes the original defect to be suppressed (ochre suppression) and involves only base-pair substitution transitions at G-C base-pairs.
Like the uvrB mutation in the Salmonella strains, the uvrA mutation causes the bacteria to be deficient in the excision of bulky lesions from the DNA, so, it is more readily mutated by certain agents (ultraviolet radiation, polycyclic hydrocarbons). Further the strain contains the pKM101 plasmid as described above for the Salmonella strains.

4.6 Bacterial cultures

The test strains of *Salmonella typhimurium* LT2 were obtained from Prof. B.N. Ames,Biochemistry Department, University of California, Berkeley, CA 94720, U.S.A.

Escherichia coli WP2uvrApKM101 was obtained from Covance Laboratories Ltd, Otley road, Harrogate, North Yorkshire HG3 1PY, England.

New batches of culture stocks frozen in 8% dimethyl sulphoxide are prepared at intervals from a central stock held in liquid nitrogen. They are regularly checked for appropriate amino acid requirement, spontaneous reversion rate, genetic characters and response to diagnostic mutagens.
Samples of each strain were grown up overnight in Nutrient broth in a 37 ± 1°C water bath with shaking. Fresh cultures were prepared before each test.

4.7 S9 mix

Composition of a 10% V/V S9 mix (final concentrations):
Co-factors:

-phosphate buffer (0.2M, pH 7.4) .....100 mM
-salts (1.65M KCl, 0.4 M $MgCl_2$).....33 and 8 mM
-glucose-6-phosphate, mono-Na salt (0.2M) ..... 5 mM
-NADP, di-Na salt (0.1M) ..... 4 mM
S9 preparation .....10% V/V

A freshly prepared solution of the co-factors was filter-sterilised by passage through a 0.2 µm membrane filter and mixed 9:1 (v/v) with freshly thawed still cold S9 preparation. This S9 mix was prepared freshly each day, and immediately used. Unused reagent was discarded.

4.8 Test material

Serial dilutions of a sterile standard solution of 5% (w/v) Amyloglucosidase PPY32789 were prepared in sterile deionised water corresponding to the final dose levels:

5000 µg - 2500 µg - 1250 µg - 625 µg - 313 µg – 156 µg substance per mL.

The dilutions were prepared freshly each day just before use.
This range of doses was applied in experiments with respectively without S9.

4.9 Top agar

0.6 % soft agar was sterilised by autoclaving.
Bottles with 200 ml melted soft agar were kept at about 55°C and added 20 ml 0.5 mM L-histidine/biotin solution for strains of *Salmonella* or 20 ml 0.5 mM tryptophan solution for *Escherichia coli*. This molten agar was divided into 2 ml aliquots in sterile glass tubes and placed in a heat-block at 45 ± 1°C.

5. Methods

5.1. Treat and plate assay

This procedure was applied with all strains.
For each assay sterile tubes were added:
- 4 ml Nutrient broth
- 4 ml S9 mix or 0.2M phosphate buffer (pH 7.4)
- 1 ml bacterial culture
- 1 ml test substance solution (6 doses) or diagnostic mutagen solution (positive control) or sterile deionised water (solvent control).

These incubation mixtures were incubated with shaking at 37 ± 1°C for 3 hours.
After incubation all bacterial suspensions were washed two times by centrifugation for 10 minutes at 2500 rpm. After the first washing the bacterial pellets were resuspended in 5 ml phosphate buffer (pH 7.4, 0.2M) and finally they were re-suspended in 1 ml phosphate buffer.
Tubes with top agar were added 0.1 ml of all washed bacterial suspensions.

5.2 Selective incubation

For each dose of the test substance and the standard mutagens three similar tubes with top agar were prepared and five tubes were prepared for the solvent control.
These tubes were poured on to minimal glucose agar plates. When the soft agar set, the plates were inverted and incubated at 37 ± 2°C for about 72 hours as described in the standard operating procedures.

After incubation the numbers of revertant colonies were counted automatically ("Cardinal" - Perceptive Instruments). Plates with less than about 20 colonies were counted manually.

5.3 Viable cell count

0.1 ml aliquots of a 10^{-6} dilution of each bacterial suspension were poured on to minimal glucose agar plates (added the required amino acids in excess) in duplicates.

5.4 Controls

The following controls were run with each experiment:

Genotype checking:
Sensitivity for crystal violet (rfa-character), (all Salmonella strains)
Sensitivity for Mitomycin C uvrB (*Salmonella*) and uvrA (*E.coli*).
Resistance to ampicillin TA98 & TA100 (*Salmonella*) and uvrA pKM101 (*E.coli*).

0.1 ml bacterial culture was spread on to nutrient agar medium. To the surface of the dried plate was added a disc of ampicillin/(Rosco Neo-Sensitabs) and two 6 mm ϕ sterile filter discs, one with 10µl 0.1% crystal violet and the other with 10µl 0.01% Mitomycin C. The plate was incubated for 48-72 hours at 37 ± 2°C.

Sterility of Amyloglucosidase standard solution and S9 mix:
0.1 ml of standard solution or S9 mix was plated on to complete medium and incubated for 48-72 hours at 37 ± 2°C. Unfortunately a temperature interval at 37 ± 1°C was noted in the protocol. This deviation has no influence on the validity of the study.

Diagnostic mutagens were used for each strain with and without S9 mix, as follows:

Mutagen	S9	Strain	µg/plate
MNNG	-	TA 1535	1.0
MNNG	-	TA 100	1.0
2-NF	-	TA 98	20.0
ICR-191	-	TA 1537	0.01
MNNG	-	WP2uvrApKM101	7.5
2-AA	+	TA 98	5.0
2-AA	+	TA 1537	5.0
2-AA	+	TA 1535	5.0
2-AA	+	TA 100	5.0
2-AA	+	WP2uvrApKM101	20.0

Deviation from protocol. In the first test series with the Salmonella strain TA1537 without S9, the positive control ICR-191 was diluted in water instead of DMSO. Since the TA1537 culture responded very well to this ICR-191 solution, this deviation has no impact on the results.

6. Results and discussion

Genetic characters

All *Salmonella* strains used in these experiments were sensitive to crystal violet and Mitomycin C. TA98 and TA100 were both resistant to ampicillin. *E.coli* WP2uvrApKM101 was sensitive to Mitomycin C. These results are as expected.

Solvent and positive controls

In general the solvent control values presented in this report are within the normal ranges experienced in our laboratory (Appendix 1) and/or the ranges reported in the literature with these strains. The one exception to this was in the second experiment with the strain TA100 in which it was slightly below what is stated in the literature. However, the experiment was accepted since culture responded very well to the positive control. Further, it should be noted that the spontaneous revertant levels are in general a bit lower in a treat and plate method than by direct plate incorporation.

All positive control chemicals induced significant increases in revertant colony numbers which fulfilled our criteria for an acceptable response (Appendix 2) and thereby confirmed the sensitivity of the test system.

Therefore all data were considered as valid.

Amyloglucosidase (Batch Number: PPY32789)

The results are represented in Table 1-10.

We consider a test substance as positive when it has induced at least a doubling in the mean number of revertant colonies per plate compared to the appropriate solvent control in one or more of the strains, in the presence or absence of S9 mix, if this response is dose related (at least 3 doses) and reproducible.
In case of a dose related and reproducible numerical increase, which is below a doubling but at least 50% higher than the solvent control, the result is considered as equivocal and needs further clarification.

Amyloglucosidase is a fluid enzyme preparation. It contains an abundance of various nutrients, and composes a rich growth medium to the test bacteria. This means, that comparison of viable counts between exposed cultures and control culture in a "treat and plate" assay reflects growth stimulation/inhibition as well as cell killing. Variation in the viable counts may cause some variation in the number of spontaneous revertant colonies.

These circumstances are reflected in the present study. No toxicity of the test substance to the bacteria is observed. On the contrary growth stimulation is evident in several test series as demonstrated by increases in the viable count of exposed cultures compared to the solvent control. Concomitantly weak numerical increases in the number of spontaneous revertant colonies are present in some test series and in two cases exceeding the 50 % criteria. This condition concerns increases in the number of spontaneous revertent colonies and should not be confused with induced revertant colonies. It is as expected and obviously without biological importance.

No treatments of any of the *Salmonella* and *E.coli* strains with Acid Metalloproteinase resulted in any increases in revertant numbers that meet the criteria for a positive or equivocal response.

7. Conclusion

The results of the bacterial mutagenicity tests described in this report give no indication of the presence of mutagenic components in this preparation of Amyloglucosidase (Batch No. PPY32789), when tested under the conditions employed in this study.

Table 1 - 10.

Table 1. Number of revertant colonies per plate obtained with ***Salmonella typhimurium*** following exposure to Amyloglucosidase (Batch No. PPY32789) in the absence of metabolic activation in the treat and plate assay.

1. experiment without S9 Mix.

Test Substance Concentration µg per mL	Number of revertants (number of colonies/plate) Base-pair substitution type							
	TA100				**TA1535**			
	Revertants		Viable cells*)		Revertants		Viable cells*)	
	Single plates	Mean	Single plates	Mean	Single plates	Mean	Single plates	Mean
5000	115 96 131	114	262 264	263	8 14 14	12	229 193	211
2500	94 117 110	107	168 152	160	11 8 9	9	189 175	182
1250	90 99 113	101	132 110	121	8 16 11	12	190 125	158
625	77 96 105	93	159 125	142	13 5 12	10	200 179	190
313	77 90 88	85	145 170	158	10 5 11	9	156 200	178
156	86 115 120	107	156 184	170	10 11 8	10	130 118	124
Solvent control	105 108 86 86 134	104	147 108	128	7 6 2 8 6	6	156 125	141
MNNG 1.0 µg	3328 3217 3109	3218	92 81	87	3421 3350 3598	3456	96 101	99

*) Number of cells x 10^7 per mL

Table 2. Number of revertant colonies per plate obtained with ***Salmonella typhimurium*** following exposure to Amyloglucosidase (Batch No. PPY32789) in the presence of metabolic activation in the treat and plate assay.

1. experiment with S9 Mix.

Test Substance Concentration µg per mL	Number of revertants (number of colonies/plate) Base-pair substition type							
	TA100				**TA1535**			
	Revertants		Viable cells*)		Revertants		Viable cells *)	
	Single plates	Mean	Single plates	Mean	Single plates	Mean	Single plates	Mean
5000	71 103 93	89	253 250	252	10 11 13	11	320 322	321
2500	123 101 80	101	210 197	204	17 15 14	15	258 310	284
1250	107 90 109	102	155 158	157	8 8 8	8	265 269	267
625	114 93 98	102	123 125	124	7 10 4	7	265 300	283
313	104 108 117	110	117 101	109	15 12 14	14	151 158	155
156	96 106 88	97	132 145	139	14 7 10	10	201 235	218
Solvent control	103 98 79 99 118	99	130 164	147	13 5 10 7 3	8	160 185	173
2-AA 5.0 µg	2096 2018 2213	2109	101 91	96	117 126 120	121	136 138	137

*) Number of cells x 10^7 per mL

Table 3. Number of revertant colonies per plate obtained with ***Salmonella typhimurium*** following exposure to Amyloglucosidase (Batch No. PPY32789) in the absence of metabolic activation in the treat and plate assay.

1. experiment without S9 Mix.

Test Substance Concentration µg per mL	Number of revertants (number of colonies/plate) Frame-shift mutation type							
	TA98				**TA1537**			
	Revertants		Viable cells*)		Revertants		Viable cells *)	
	Single plates	Mean	Single plates	Mean	Single plates	Mean	Single plates	Mean
5000	16 13 15	15	331 305	318	13 12 10	12	228 240	234
2500	15 8 17	13	254 261	258	9 2 7	6	193 211	202
1250	18 20 19	19	218 207	213	6 12 6	8	179 183	181
625	13 13 23	16	163 175	169	5 7 9	7	217 155	186
313	12 17 10	13	188 190	189	10 8 5	8	155 144	150
156	11 24 9	15	178 209	194	7 10 9	9	156 146	151
Solvent control	13 24 23 17 15	18	180 179	180	8 6 5 3 9	6	218 158	188
2-NF 20.0 µg	1289 1158 1256	1234	133 157	145				
ICR-191 0.01 µg					1873 ¤) 1964 ¤) 2096 ¤)	1978	183 194	189

*) Number of cells x 10^7 per mL
¤) See section 5.4: Deviation from protocol

Table 4. Number of revertant colonies per plate obtained with ***Salmonella typhimurium*** following exposure to Amyloglucosidase (Batch No. PPY32789) in the presence of metabolic activation in the treat and plate assay.

1. experiment with S9 Mix.

Test Substance Concentration µg per mL	Number of revertants (number of colonies/plate) Frame-shift mutation type							
	TA98				**TA1537**			
	Revertants		Viable cells*)		Revertants		Viable cells *)	
	Single plates	Mean	Single plates	Mean	Single plates	Mean	Single plates	Mean
5000	26 26 26	26	206 238	222	16 10 9	12	137 147	142
2500	17 28 28	24	175 183	179	7 11 9	9	123 114	119
1250	23 17 20	20	157 129	143	18 14 17	16	126 130	128
625	17 20 18	18	148 157	153	7 12 10	10	123 105	114
313	21 22 20	21	206 223	215	18 18 13	16	124 92	108
156	22 16 16	18	182 178	180	16 12 17	15	115 119	117
Solvent control	23 22 22 22 21	22	173 200	187	15 13 16 17 13	15	112 108	110
2-AA 5.0 µg	2344 2355 2384	2361	117 143	130	260 239 250	250	86 108	97

*) Number of cells x 10^7 per mL

Table 5. Number of revertant colonies per plate obtained with ***E.coli*** **WP2uvrApKM101** following exposure to Amyloglucosidase (Batch No. PPY32789) in the absence and presence of metabolic activation in the "treat and plate" assay.

1. experiment without and with S9 Mix.

Test Substance Concentration µg per mL	Number of revertants (number of colonies/plate) Base-pair substition type							
	Without S9				**With S9**			
	Revertants		Viable cells*)		Revertants		Viable cells *)	
	Single plates	Mean	Single plates	Mean	Single plates	Mean	Single plates	Mean
5000	231 202 197	210	362 352	357	299 291 236	275	510 510	510
2500	191 189 209	196	308 294	301	276 238 227	247	428 393	411
1250	193 195 205	198	254 216	235	221 237 223	227	378 346	362
625	183 148 141	157	194 202	198	209 200 188	199	347 321	334
313	182 173 172	176	234 216	225	212 238 202	217	323 329	326
156	183 188 188	186	205 193	199	213 211 218	214	335 339	337
Solvent control	183 172 169 140 173	167	179 205	192	216 226 217 183 215	211	364 357	361
MNNG 7,5µg	953 1035 969	986	151 130	141				
2-AA 20,0 µg					1561 1380 1485	1475	274 236	255

*) Number of cells x 10^7 per mL

Table 6. Number of revertant colonies per plate obtained with ***Salmonella typhimurium*** following exposure to Amyloglucosidase (Batch No. PPY32789) in the absence of metabolic activation in the treat and plate assay.

2. experiment without S9 Mix.

Test Substance Concentration µg per mL	Number of revertants (number of colonies/plate) Base-pair substition type							
	TA100				**TA1535**			
	Revertants		Viable cells*)		Revertants		Viable cells *)	
	Single plates	Mean	Single plates	Mean	Single plates	Mean	Single plates	Mean
5000	67 105 117	96	272 215	244	12 7 13	11	271 301	286
2500	66 103 71	80	159 166	163	11 20 14	15	202 159	181
1250	80 80 101	87	132 123	128	10 9 9	9	174 166	170
625	66 87 61	71	120 114	117	13 5 6	8	119 134	127
313	75 70 82	76	130 126	128	9 14 5	9	130 117	124
156	69 79 67	72	115 164	140	7 8 8	8	170 162	166
Solvent control	53 86 69 60 80	70	92 124	108	11 10 7 8 9	9	131 137	134
MNNG 1.0 µg	4029 3861 3707	3866	124 110	117	3253 3445 3050	3249	148 125	137

*) Number of cells x 10^7 per mL

Table 7. Number of revertant colonies per plate obtained with ***Salmonella typhimurium*** following exposure to Amyloglucosidase (Batch No. PPY32789) in the presence of metabolic activation in the treat and plate assay.

2. Experiment with S9 Mix.

Test Substance Concentration µg per mL	Number of revertants (number of colonies/plate) Base-pair substation type							
	TA100				**TA1535**			
	Revertants		Viable cells*)		Revertants		Viable cells*)	
	Single plates	Mean	Single plates	Mean	Single plates	Mean	Single plates	Mean
5000	98 86 83	89	303 272	288	8 6 6	7	350 309	330
2500	90 93 104	96	194 221	208	12 14 9	12	269 249	259
1250	93 91 99	94	215 231	223	11 7 12	10	276 320	298
625	96 94 106	99	168 177	173	11 6 10	9	245 250	248
313	86 120 83	96	200 196	198	7 11 10	9	189 200	195
156	77 107 97	94	160 172	166	6 13 8	9	193 183	188
Solvent control	108 92 82 87 109	96	113 108	111	11 8 10 9 11	10	125 114	120
2-AA 5.0 µg	2539 2442 2297	2426	130 114	122	182 207 169	186	87 61	74

*) Number of cells x 10^7 per mL

Table 8. Number of revertant colonies per plate obtained with ***Salmonella typhimurium*** following exposure to Amyloglucosidase (Batch No. PPY32789) in the absence of metabolic activation in the treat and plate assay.

2. experiment without S9 Mix.

Test Substance Concentration µg per mL	Number of revertants (number of colonies/plate) Frame-shift mutation type							
	TA98				**TA1537**			
	Revertants		Viable cells*)		Revertants		Viable cells *)	
	Single plates	Mean	Single plates	Mean	Single plates	Mean	Single plates	Mean
5000	28 22 29	26	158 69	114	6 9 10	8	269 240	255
2500	29 16 29	25	126 134	130	7 11 9	9	211 228	220
1250	25 20 16	20	157 141	149	6 10 9	8	236 304	270
625	21 26 37	28	238 275	257	7 7 16	10	266 298	282
313	25 25 26	25	91 56	74	8 8 6	7	166 156	161
156	15 28 28	24	161 174	168	8 10 10	9	196 211	204
Solvent control	20 21 18 34 31	25	115 128	122	15 8 15 11 8	11	204 150	177
2-NF 20.0 µg	693 634 612	646	150 98	124				
ICR-191 0,01 µg					1484 1429 1457	1457	145 156	151

*) Number of cells x 10^7 per mL

Table 9. Number of revertant colonies per plate obtained with ***Salmonella typhimurium*** following exposure to Amyloglucosidase (Batch No. PPY32789) in the presence of metabolic activation in the treat and plate assay.

2. experiment with S9 Mix.

Test Substance Concentration µg per mL	Number of revertants (number of colonies/plate) Frame-shift mutation type							
	TA98				**TA1537**			
	Revertants		Viable cells*)		Revertants		Viable cells *)	
	Single plates	Mean	Single plates	Mean	Single plates	Mean	Single plates	Mean
5000	34 31 31	32	283 260	272	11 10 12	11	177 179	178
2500	37 36 37	37	253 224	239	16 7 16	13	150 204	177
1250	23 33 26	27	286 238	262	18 10 15	14	175 175	175
625	21 31 25	26	197 260	229	9 10 8	9	155 151	153
313	27 37 31	32	216 232	224	11 11 13	12	153 152	153
156	36 29 34	33	225 281	253	13 14 20	16	163 163	163
Solvent control	38 39 37 20 25	32	205 227	216	13 13 12 13 17	14	158 129	144
2-AA 5.0 µg	2329 2008 2031	2123	174 180	177	193 159 163	172	91 114	103

*) Number of cells x 10^7 per mL

Table 10. Number of revertant colonies per plate obtained with ***E.coli*** **WP2uvrApKM101** following exposure to Amyloglucosidase (Batch No. PPY32789) in the absence and presence of metabolic activation in the "treat and plate" assay.

2. experiment without and with S9 Mix.

Test Substance Concentration µg per mL	Number of revertants (number of colonies/plate) Base-pair substition type							
	Without S9				**With S9**			
	Revertants		Viable cells*)		Revertants		Viable cells *)	
	Single plates	Mean	Single plates	Mean	Single plates	Mean	Single plates	Mean
5000	341 355 271	322	397 378	388	367 288 255	303	579 545	562
2500	229 223 218	223	303 298	301	259 265 231	252	537 488	513
1250	250 228 231	236	229 330	280	304 222 248	258	456 426	441
625	194 189 175	186	286 252	269	217 177 197	197	311 375	343
313	234 202 182	206	258 258	258	218 205 220	214	340 369	355
156	220 168 194	194	216 217	217	254 206 239	233	442 471	457
Solvent control	173 177 173 210 186	184	234 225	230	250 251 193 215 200	222	367 358	363
MNNG 7,5 µg	1952 1987 2165	2035	169 206	188				
2-AA 20,0 µg					1445 1360 1316	1374	280 272	276

*) Number of cells x 10^7 per mL

Appendix 1

Historical control data

Negative control (purified water) for *S. typhimurium* strains and *E.coli* WP2uvrA pKM101 in the treat and plate assay. (SOP: TOX-SM-1006 and TOX-SM-1007)

Strain	S9	Number of determinations	Mean number of revertants per plate	SD	Range *)	
					lower	upper
TA1535	÷	12	8	2	4	13
	+	13	10	3	7	17
TA100	÷	11	96	11	85	117
	+	13	117	15	89	140
TA1537	÷	12	7	3	4	14
	+	14	11	2	7	15
TA98	÷	12	22	5	16	33
	+	11	30	8	20	41
WP2 uvrA pKM101	÷	13	190	21	164	241
	+	13	200	21	174	241

The above are pooled data from a number of independent determinations selected from studies conducted over the period September 2010 to September 2011. Only determinations, which were obviously vitiated by errors, have been omitted.

*) Ranges stated are the maximum and minimum mean revertant colony counts from the data sets sampled.

Appendix 2

Historical control data

Positive control ranges for *S. typhimurium* strains and *E.coli* WP2uvrA pKM101 in the treat and plate assay. (SOP: TOX-SM-1006 and TOX-SM-1007).

Strain	S9	Number of determinations	Chemical	Mean number of revertants per plate	SD	Range *)	
						lower	upper
TA1535	÷	12	MNNG 1 µg/plate	4322	1175	1849	5446
	+	13	2-AA 5 µg/plate	167	19	140	209
TA100	÷	11	MNNG 1 µg/plate	3897	714	2570	4639
	+	13	2-AA 5 µg/plate	2063	428	1241	2681
TA1537	÷	12	ICR-191 0.01µg/plate	1669	228	1260	2071
	+	14	2-AA 5 µg/plate	158	55	92	310
TA98	÷	12	2-NF 20 µg/plate	1002	221	648	1415
	+	11	2-AA 5 µg/plate	1963	515	1116	2972
WP2 uvrA pKM101	÷	13	MNNG 7.5 µg/plate	1191	272	839	1728
	+	13	2-AA 20 µg/plate	1426	260	878	1741

The above are pooled data from a number of independent determinations selected from studies conducted over the period period September 2010 to September 2011. Only determinations, which were obviously vitiated by errors, have been omitted.

*) Ranges stated are the maximum and minimum mean revertant colony counts from the data sets sampled.

Appendix 3

PREPARATION OF MEDIA

1. **Top-agar - histidine-deficient soft agar**

Agar, Merck	0.6 g
NaCl	0.5 g
Distilled water to	100 ml

The medium was autoclaved for 15 minutes at 121°C. After cooling to about 60°C, 10 ml of a sterile aqueous solution of 0.5 mM biotin - 0.5 mM histidine was added aseptically.

2. **Nutrient broth - histidine-rich broth**

Difco nutrient broth	8 g
NaCl	5 g
Distilled water to	1 litre

The medium was autoclaved for 15 minutes at 121°C.

3. **Nutrient agar - histidine-rich agar medium**

Agar, Merck	15 g
Oxoid nutrient broth No. 2	25 g
Distilled water to	1 litre

The medium was autoclaved for 15 minutes at 121°C.

4. **Minimal medium**

This was Vogel-Bonner minimal "E" medium with 2% glucose, prepared as follows:

Solution A (Vogel-Bonner medium E, 20X)

$MgSO_4$ $7H_2O$	4 g
Citric acid, monohydrate	40 g
K_2HPO_4	200 g
$NaHNH_4$ $4H_2O$	70 g
Distilled water to	1000 ml

The solution was sterilized by filtration.

Solution B (40% glucose)

Glucose	40 g
Distilled water to	100 ml

This solution was sterilized by filtration.

Solution C (Agar base)

Agar, Merck	16.7 g
Distilled water to	1000 ml

Solution C was autoclaved for 15 minutes at 121°C. After cooling to 60°C, 450 ml of solution C was aseptically added 25 ml solution A and 25 ml solution B.

Final Report

Study Title	Induction of micronuclei in cultured human peripheral blood lymphocytes
Test Article	Amyloglycosidase PPY 32789
Author	J Whitwell BSc
Sponsor	Novozymes A/S Toxicology Krogshoejsvej 36 2880 Bagsvaerd Denmark
Study Monitor	P Pedersen
Test Facility	Covance Laboratories Ltd Otley Road, Harrogate North Yorkshire HG3 1PY, ENGLAND
Covance Client Identifier	1002135
Covance Study Number	8259272
Novozymes Reference number	NZ20126003
Report Issued	June 2012
Page Number	1 of 41

COVANCE
THE DEVELOPMENT SERVICES COMPANY

STUDY DIRECTOR AUTHENTICATION AND GLP COMPLIANCE STATEMENT

Amyloglycosidase PPY 32789: Induction of micronuclei in cultured human peripheral blood lymphocytes

I, the undersigned, hereby declare that the work was performed under my supervision and that the findings provide a true and accurate record of the results obtained.

The study was performed in accordance with the agreed protocol and with Covance Laboratories Limited Standard Operating Procedures, unless otherwise stated, and the study objectives were achieved.

The study was conducted in compliance with the United Kingdom Good Laboratory Practice Regulations 1999, Statutory Instrument No. 3106 as amended by the Good Laboratory Practice (Codification Amendments Etc.) Regulations 2004 and the OECD Principles on Good Laboratory Practice (revised 1997, issued January 1998) ENV/MC/CHEM (98) 17.

13 June 2012
Date

Study Director

QUALITY ASSURANCE STATEMENT

Amyloglycosidase PPY 32789: Induction of micronuclei in cultured human peripheral blood lymphocytes

This study has been reviewed by the GLP Quality Assurance Unit of Covance and the report accurately reflects the raw data. The following inspections were conducted and findings reported to the Study Director (SD) and associated management.

Critical procedures, which are performed routinely in an operational area, may be audited as part of a "process" inspection programme. This can be in addition to phases scheduled on an individual study basis. Selected process inspections conducted and considered applicable to this study are included below.

In addition to the inspection programmes detailed below, a facility inspection programme is also operated. Details of this programme, which covers all areas of the facility annually (at a minimum), are set out in standard operating procedures.

Inspection Dates		Phase	Date Reported to SD and SD Management
From	To		
22 Dec 2011	22 Dec 2011	Protocol Review	22 Dec 2011
11 Apr 2012	12 Apr 2012	Draft Report and Data Review	12 Apr 2012
13 Jun 2012	13 Jun 2012	Final Report Review	13 Jun 2012

Process

Inspection Dates		Phase	Date Reported to SD and SD Management
From	To		
22 Dec 2011	22 Dec 2011	Stock Solution Preparation	22 Dec 2011
09 Jan 2012	09 Jan 2012	Culture Establishment	09 Jan 2012
17 Jan 2012	17 Jan 2012	Dose Preparation	17 Jan 2012
27 Jan 2012	27 Jan 2012	Outworker Documents	30 Jan 2012
14 Feb 2012	14 Feb 2012	Dose Preparation	15 Feb 2012
06 Mar 2012	06 Mar 2012	Dose Preparation	06 Mar 2012

13/06/12

Date

Quality Assurance Unit

REVIEWING SCIENTIST'S STATEMENT

Amyloglycosidase PPY 32789: Induction of micronuclei in cultured human peripheral blood lymphocytes

I, the undersigned, hereby declare that I have reviewed this report in conjunction with the Study Director and that the interpretation and presentation of the data in the report are consistent with the results obtained.

_____ 28 May 2012
Date

Scientist

RESPONSIBLE PERSONNEL

Amyloglycosidase PPY 32789: Induction of micronuclei in cultured human peripheral blood lymphocytes

The following personnel were responsible for key elements of the study:

Study Director	J Whitwell
Laboratory Supervisor	K Jenner
Study Monitor [1]	P Pedersen

[1] Located at Novozymes A/S, Denmark.

ARCHIVE STATEMENT

Amyloglycosidase PPY 32789: Induction of micronuclei in cultured human peripheral blood lymphocytes

All primary data, or authenticated copies thereof, specimens and the final report will be retained in the Covance Laboratories Limited archives for five years after issue of the final report. At the end of the specified archive period the Sponsor will be contacted to determine whether the data should be returned, retained or destroyed on their behalf. Sponsors will be notified of the financial implications of each of these options at that time. One copy of the protocol and final report will be held in the Covance Laboratories Limited archives as per Covance company policy.

CONTENTS

SUMMARY

Amyloglycosidase PPY 32789 was tested in an *in vitro* micronucleus assay using duplicate human lymphocyte cultures prepared from the pooled blood of two male donors in a single experiment. Treatments covering a broad range of concentrations, separated by narrow intervals, were performed both in the absence and presence of metabolic activation (S-9) from Aroclor 1254 induced animals. The test article was formulated in water for irrigation (purified water) and the highest concentration tested in the Micronucleus Experiment, 5000 µg/mL (an acceptable maximum concentration for *in vitro* chromosome aberration studies according to current regulatory guidelines), was determined following a preliminary cytotoxicity Range-Finder Experiment.

Treatments were conducted (as detailed in the following summary table) 48 hours following mitogen stimulation by phytohaemagglutinin (PHA). The test article concentrations for micronucleus analysis were selected by evaluating the effect of Amyloglycosidase PPY 32789 on the replication index (RI). In the Micronucleus Experiment, micronuclei were analysed at three concentrations and a summary of the micronucleus data is presented in Table 1:

Table 1: Micronucleus Experiment (48 hour PHA) – Results summary

Treatment	Concentration (µg/mL)	Cytotoxicity (%)	Mean MNBN cell frequency (%)	Historical Control Range (%) #	Statistical significance
3+21 hour -S-9	Vehicle [a]	-	0.33	0.20 - 0.80	-
	3000	0	0.40		NS
	4000	0	0.30		NS
	5000	0	0.50		NS
	*MMC, 0.80	ND	13.15		$p \leq 0.001$
3+21 hour +S-9	Vehicle [a]	-	0.15	0.10 – 1.10	-
	3000	2	0.40		$p \leq 0.05$
	4000	5	0.35		NS
	5000	4	0.20		NS
	*CPA, 12.5	ND	3.50		$p \leq 0.001$
24+24 hour -S-9	Vehicle [a]	-	0.25	0.00 – 1.10**	-
	3000	0	0.30		NS
	4000	0	0.60		$p \leq 0.05$
	5000	0	0.10		NS
	*VIN, 0.08	ND	7.95		$p \leq 0.001$

[a] Vehicle control was purified water.
* Positive control
#95th percentile of the observed range
NS = Not significant / ND = Not determined
** Historical control range was calculated on 24+0 hour –S-9 treatments. Range included as a guide only.

Appropriate negative (vehicle) control cultures were included in the test system under each treatment condition. The proportion of micronucleated binucleate cells (MNBN) in these cultures fell within current historical vehicle control (normal) ranges. Mitomycin C (MMC) and Vinblastine (VIN) were employed as clastogenic and aneugenic positive control chemicals respectively in the absence of rat liver S-9. Cyclophosphamide (CPA) was employed as a clastogenic positive control chemical in the presence of rat liver S-9. Cells receiving these were sampled in the Micronucleus Experiment at 24 hours (CPA, MMC) or 48 hours (VIN) after the start of treatment; all compounds induced statistically significant increases in the proportion of cells with micronuclei.

All acceptance criteria were met and the study was therefore considered as valid.

Pulse 3+21 hour treatment of cells with Amyloglycosidase PPY 32789 in the absence and presence of S-9 resulted in frequencies of MNBN cells which were similar to and not significantly ($p \leq 0.05$) higher than those observed in concurrent vehicle controls for the majority of concentrations analysed. The single exception to this was for the

lowest concentration analysed following treatment in the presence of S-9 (3000 µg/mL) where a small increase was noted (mean MNBN cell frequency = 0.4% compared to 0.15% in the concurrent vehicle control). However, as the MNBN cell frequency of both replicate cultures at this and all other Amyloglycosidase PPY 32789 treated cultures fell within historical vehicle control (normal) values, this isolated statistical increase was not considered of biological importance.

Following 24+24 hour –S-9 treatment a further small but statistically significant increase in MNBN cells was noted at the intermediate concentration analysed (4000 µg/mL). However, there were no instances of MNBN cell values exceeding normal values for any test article treated culture (at this or any other concentration), with higher and lower concentrations analysed showing mean MNBN cell values similar to those observed for the vehicle control. As such, this isolated statistical increase was not considered of biological importance.

It is concluded that Amyloglycosidase PPY 32789 did not induce micronuclei in cultured human peripheral blood lymphocytes following treatment in the absence and presence of a rat liver metabolic activation system (S-9). Concentrations were tested up to 5000 µg/mL, a recommended regulatory maximum concentration for *in vitro* cytogenetic assays.

INTRODUCTION

Chromosome defects are recognised as the basis of a number of human genetic diseases (Mitelman, 1991). Assays for the detection of chromosome damage in mammalian cells *in vitro* are recommended in regulatory guidelines as a complement to Ames tests in a genotoxicity test battery. There is a large database on the use of chromosomal assays for screening purposes (Preston *et al.*, 1981; Fenech, 1998; Fenech *et al.*, 2003). The use of human peripheral blood lymphocytes is recommended because the cells are only used in short-term culture and maintain a stable karyotype (Evans & O'Riordan, 1975). Experiments with these cells can also be performed in conjunction with a rat liver metabolising system (S-9) since, for short incubation periods, no toxicity is induced by the liver homogenate itself.

An alternative to measuring structural aberrations in mitotic cells is to measure micronuclei. These are produced from whole chromosomes or acentric fragments that are unable to attach to the spindle at mitosis and appear during the next interphase as small darkly staining bodies adjacent to the main daughter nucleus. Cytochalasin B (Cyto-B), if added to cultures, inhibits cytokinesis (cell division) but not karyokinesis (nuclear division) resulting in the formation of binucleate cells (Fenech & Morley, 1985). If micronuclei are counted in binucleate cells, then a measurement of micronucleus induction resulting from cell division can be obtained.

Theoretical considerations, together with published data (Lorge *et al.*, 2006), indicate that most aneugens and clastogens will be detected by a short term treatment period of 3-6 hours in the presence and absence of S-9 followed by removal of the test article and a growth period of 1.5-2.0 cell cycles (Fenech & Morley, 1986).

The most efficient approach is to test lymphocytes 44-48 hours post-mitogen stimulation by PHA, when cycle synchronisation will have dissipated (Fenech, 2007).

The test article was added at 48 hours following culture initiation (stimulation by PHA). Cells were exposed to the test article for 3 hours in the absence and presence of S-9 (from rats induced with Aroclor 1254). These cultures were sampled 24 hours after the beginning of treatment (i.e. 72 hours after culture initiation).

In addition, an extended 24 hour treatment (equivalent to approximately 1.5 to 2 times the average generation time of cultured lymphocytes from the panel of donors used in this laboratory) with 24 hour recovery in the absence of S-9 was included. These

cultures were sampled 48 hours after the beginning of treatment (i.e. 96 hours after culture initiation).

The objective of this study was to evaluate the clastogenic and aneugenic potential of Amyloglycosidase PPY 32789 by examining its effects on the frequency of micronuclei in cultured human peripheral blood lymphocytes treated in the absence and presence of S-9.

The test methodology is based on OECD guideline 487 (OECD, 2010) and accepted scientific/regulatory principles described in current guidelines for clastogenicity testing *in vitro* (Fenech, 1998; Fenech *et al.*, 2003; Rosefort *et al.*, 2004; Elhajouji *et al.*, 1998; Migliore & Nieri, 1991; Galloway *et al.*, 1994; Aardema *et al.*, 1998; Miller *et al.*, 1998; ICH-S2A, 1995; ICH-S2B, 1997; Fenech *et al.*, 1999; Thybaud *et al.*, 2007).

This study was performed according to the protocol.

The study was initiated on 20 December 2011. Experimental work started on 30 January 2012 and was completed on 12 March 2012. The study completion date is considered to be the date the Study Director signs the final report.

MATERIALS

Test article

Amyloglycosidase PPY 32789, batch number PPY 32789, was a frozen brown liquid (nominal volume stated as 100 mL). It was received on 23 December 2012. Following receipt the test article was thawed and sub-divided into three aliquots (1 x 15 mL for use on the Range-Finder Experiment and 2 x 42 mL aliquots for use on the Micronucleus Experiment). These aliquots were then re-frozen and stored at -20°C nominal in the dark in order that the test article should be thawed and re-frozen on no more than one occasion.

Purity / activity was stated as 392 AGU/g, but for the purposes of this study was considered as 100%. The expiry date was given as 21 August 2021. The Documentation of Test Material, provided by the Sponsor, is presented in Appendix 5. The test article information and Documentation of Test Material provided by the Sponsor are considered an adequate description of the characterisation, purity and stability of the test article. Determinations of stability and characteristics of the test article were the responsibility of the Sponsor.

Amyloglycosidase PPY 32789 is a high molecular weight protein, which was formulated in water for irrigation (purified water) to a concentration of 50 mg/mL (weighed out as received), equivalent to 5000 µg/mL final culture concentration. Once de-frosted, the test article was stirred gently (under magnetic stirring conditions) both prior to and following weighing.

No preliminary solubility trials were conducted as part of this study.

Test article stock solutions were prepared by formulating Amyloglycosidase PPY 32789 under subdued light in purified water, with the aid of stirring to give 50 mg/mL. The stock solutions were membrane filter-sterilised (Pall Acrodisc 32, 0.2 µm pore size) and subsequent dilutions made using purified water. The test article solutions were protected from light and used within approximately 2.5 hours of removal of the test article from refrigerated conditions as shown in Table 2:

Table 2: Amyloglycosidase PPY 32789 Concentration Ranges Tested

Experiment	Treatment	Concentration range (mg/mL)	Final concentration range (µg/mL)
Range-Finder	3+21, -S-9	0.1814 to 50.00	18.14 to 5000
	3+21, +S-9	0.1814 to 50.00	18.14 to 5000
	24+24, -S-9	0.1814 to 50.00	18.14 to 5000
Micronucleus Experiment	3+21, -S-9	10.00 to 50.00	1000 to 5000
	3+21, +S-9	10.00 to 50.00	1000 to 5000
	24+24, -S-9	5.000 to 50.00	500.0 to 5000

Controls

Sterile purified water was added to cultures designated as negative (vehicle) controls as described in the methods section of this report.

Table 3: Positive Controls

Chemical	Stock concentration (mg/mL)*	Final concentration (µg/mL)	S-9
Mitomycin C	0.060	0.60	-
(MMC) **	0.080	0.80	-
Cyclophosphamide	0.625	6.25	+
(CPA) ***	1.25	12.50	+
Vinblastine	0.008	0.08	-
(VIN) **	0.010	0.10	-
	0.012	0.12	-

* In the Micronucleus Experiment, CPA was dissolved in anhydrous analytical grade dimethyl sulphoxide (DMSO), frozen (<-50°C) and thawed immediately prior to use or prepared immediately prior to use and diluted accordingly. VIN and MMC were dissolved in purified water immediately prior to use.

** Obtained from Sigma-Aldrich Chemical Co, Poole, UK.

*** Obtained from Acros Organics, Loughborough, Leicestershire, UK.

For the 3+21 hour treatments in the absence of S-9, MMC was used as the positive control. For the 24+24 hour treatments, VIN was used as the positive control.

Metabolic activation system

The mammalian liver post-mitochondrial fraction (S-9) used for metabolic activation was obtained from Molecular Toxicology Incorporated, USA where it is prepared from male Sprague Dawley rats induced with Aroclor 1254. The batches of MolTox™ S-9 were stored frozen in aliquots at <-50°C prior to use. Each batch was checked by the manufacturer for sterility, protein content, ability to convert known promutagens to bacterial mutagens and cytochrome P-450-catalyzed enzyme activities (alkoxyresorufin-O-dealkylase activities). The quality control statements, relating to the batches of S-9 preparation used, are included in Appendix 4 of this report.

The S-9 mix was prepared in the following way:

Glucose-6-phosphate (G6P: 180 mg/mL), β-Nicotinamide adenine dinucleotide phosphate (NADP: 25 mg/mL), Potassium chloride (KCl: 150 mM) and rat liver S-9 were mixed in the ratio 1:1:1:2. For all cultures treated in the presence of S-9, an aliquot of the mix was added to each cell culture to achieve the required final concentration of test article in a total of 10 mL. The final concentration of the liver homogenate in the test system was 2%.

Cultures treated in the absence of S-9 received an equivalent volume of KCl (150 mM).

Blood cultures

Blood from two healthy, non-smoking male volunteers from a panel of donors at Covance was used for each experiment in this study:

Table 4: Blood Cultures

Experiment	Donor Sex	Donor Age (years)	Donor Identity
Range-Finder	Male	26, 34	9817, 8844
Micronucleus Experiment	Male	31, 26	6747, 9817

No volunteer was suspected of any virus infection or exposed to high levels of radiation or hazardous chemicals. All volunteers are non-smokers and are not heavy drinkers of alcohol. Donors were not taking any form of medication. The measured cell cycle time of the donors used at Covance falls within the range 13 +/- 2 hours. For each experiment, an appropriate volume of whole blood was drawn from the peripheral circulation into heparinised tubes within two days of culture initiation. Blood was stored refrigerated and pooled using equal volumes from each donor prior to use.

Whole blood cultures were established in sterile disposable centrifuge tubes by placing 0.4 mL of pooled heparinised blood into 8.1 mL HEPES-buffered RPMI medium containing 10% (v/v) heat inactivated foetal calf serum and 0.52% penicillin / streptomycin, so that the final volume following addition of S-9 mix/KCl and the test article in its chosen vehicle was 10 mL. The mitogen Phytohaemagglutinin (PHA, reagent grade) was included in the culture medium at a concentration of approximately 2% of culture to stimulate the lymphocytes to divide. Blood cultures were incubated at 37 ± 1°C for 48 hours and rocked continuously.

METHODS

The test system was suitably labelled (using a colour-coded procedure) to clearly identify the study number, assay type, experiment number, treatment time, sex of the donor, test article concentration (if applicable), positive and negative controls.

Cytotoxicity Range-Finder

S-9 mix or KCl (0.5 mL/culture) was added appropriately. Cultures were treated with the test article or vehicle control (1.0 mL/culture) as indicated in Table 5. Positive control treatments were not included.

The final culture volume was 10 mL.

Cultures were incubated at 37 ± 1°C for the designated exposure time.

Micronucleus Experiment

Immediately prior to treatment, all positive control cultures had 0.9 mL culture medium added to give a final pre-treatment volume of 9.4 mL.

S-9 mix or KCl (0.5 mL/culture) was added appropriately. Cultures were treated with the test article or vehicle controls (1.0 mL/culture) or positive control cultures (0.1 mL/culture) as indicated in Table 5. The final culture volume was 10 mL. Cultures were incubated at 37 ± 1°C for the designated exposure time.

This scheme is illustrated as follows:

Table 5: Treatment Scheme

Treatment	S-9	Number of cultures			
		Cytotoxicity Range-Finder		Micronucleus Experiment	
		3+21*	24+24*	3+21*	24+24*
Negative control	-	2	2	4	4
	+	2		4	
Test article	-	1	1	2	2
	+	1		2	
Positive controls	-			2	2
	+			2	

* Hours treatment + hours recovery

For removal of the test article, cells were pelleted (approximately 300 *g*, 10 minutes), washed twice with sterile saline (pre-warmed in an incubator set to 37 ± 1°C), and resuspended in fresh pre-warmed medium containing foetal calf serum and penicillin / streptomycin. At the appropriate times Cytochalasin B formulated in DMSO was added to post wash-off culture medium to give a final concentration of 6 µg/mL per culture.

Table 6: Summary of treatment conditions

Duration of treatment (hours)	S-9	Hours after culture initiation			
		Addition of test article	Removal of test article	Addition of Cytochalasin B	Harvest time
3	-	48	51	51*	72
24	-	48	72	72*	96
3	+	48	51	51*	72

* Approximate times

Changes in osmolality of more than 50 mOsm/kg and fluctuations in pH of more than one unit may be responsible for an increase in chromosome aberrations (Scott *et al.*, 1991; Brusick, 1986). Osmolality and pH measurements on post-treatment incubation medium were taken in the cytotoxicity Range-Finder Experiment.

Harvesting

At the defined sampling time, cultures were centrifuged at approximately 300 *g* for 10 minutes, the supernatant removed and discarded and cells resuspended in 4 mL (hypotonic) 0.075 M KCl at 37 ± 1°C for 4 minutes to allow cell swelling to occur. Cells were then fixed by dropping the KCl suspension into fresh, cold methanol/glacial acetic acid (3:1, v/v). The fixative was changed by centrifugation (approximately 300 *g*, 10 minutes) and resuspension. This procedure was repeated as necessary (centrifuging at approximately 1250 *g*, 2-3 minutes) until the cell pellets were clean.

Slide preparation

Lymphocytes were kept in fixative at 2-8°C prior to slide preparation for a minimum of 3 hours to ensure that cells were adequately fixed. Cells were centrifuged (approximately 1250 *g*, two to three minutes) and resuspended in a minimal amount of fresh fixative (if required) to give a milky suspension. Several drops of cell suspension were gently spread onto multiple clean, dry microscope slides. Slides were

air-dried then stored protected from light at room temperature prior to staining. Slides were stained by immersion in 125 µg/mL Acridine Orange in phosphate buffered saline (PBS), pH 6.8 for approximately 10 seconds, washed with PBS (with agitation) for a few seconds before transfer and immersion in a second container of PBS for approximately 8-10 minutes. Slides were air-dried and stored protected from light at room temperature prior to analysis.

Selection of concentrations for Micronucleus Experiment

Slides from the cytotoxicity Range-Finder Experiment were examined, uncoded, for proportions of mono-, bi- and multinucleate cells, to a minimum of 200 cells per concentration. From these data the replication index (RI) was determined.

The Replication Index (RI), which indicates the relative number of nuclei compared to controls was determined using the formulae below:

$$RI = \frac{\textit{number binucleate cells} + 2(\textit{number multinucleate cells})}{\textit{total number of cells in treated cultures}}$$

Relative RI (expressed in terms of percentage) for each treated culture was calculated as follows:

$$\textit{Relative RI (\%)} = \frac{\textit{RI of treated cultures}}{\textit{RI of vehicle controls}} \quad x100$$

Cytotoxicity (%) is expressed as (100 – Relative RI).

A selection of random fields was observed from enough treatments to determine whether chemically induced cell cycle delay or cytotoxicity has occurred.

A suitable range of concentrations was selected for the Micronucleus Experiment based on these toxicity data.

Selection of concentrations for micronucleus analysis (Micronucleus Experiment only)

Slides were examined, uncoded, for proportions of mono-, bi- and multinucleate cells to a minimum of 500 cells per culture.

The highest concentration for micronucleus analysis was to be one at which approximately 55% (typically 50-60%) reduction in RI had occurred, or the highest concentration tested.

For each treatment regime, two vehicle control cultures ('A' and 'B') were initially analysed for micronuclei. However, in this study to aid data interpretation additional micronucleus analysis was performed on the vehicle control 'C' and 'D' cultures.

Positive control concentration, which gave satisfactory responses in terms of quality and quantity of binucleated cells and numbers of micronuclei, were analysed.

Slide analysis

Slides from the CPA, MMC and VIN positive control treatments were checked to ensure that the system had operated satisfactorily. All slides for analysis were coded, using randomly generated letters, by an individual not connected with the scoring of the slides. Labels with only the study number, assay type, experiment number, the sex of the donor and the code were used to cover treatment details on the slides.

Immediately prior to analysis 1-2 drops of PBS were added to the slides before mounting with glass coverslips. One thousand binucleate cells from each culture (2000 per test article concentration, 4000 per vehicle) were analysed for micronuclei. The number of cells containing micronuclei and the number of micronuclei per cell on each slide was noted. Observations were recorded on raw data sheets. The microscope stage co-ordinates of the first six micronucleated cells were recorded.

In order to clarify the data from the 3+21 hour –S-9 treatment, a further 1000 BN cells was subsequently analysed from each replicate culture (A' and 'B') from concentration 5000 μg/mL.

Binucleate cells were only included in the analysis if all of the following criteria were met:

1. The cytoplasm remained essentially intact, and

2. The daughter nuclei were of approximately equal size.

A micronucleus was only recorded if it met the following criteria:

1. The micronucleus had the same staining characteristics and a similar morphology to the main nuclei, and

2. Any micronucleus present was separate in the cytoplasm or only just touching a main nucleus, and

3. Micronuclei were smooth edged and smaller than approximately one third the diameter of the main nuclei.

Micronucleus analysis was not conducted on slides generated from the Range-Finder treatments.

Slide analysis was performed by competent analysts trained in the applicable Covance Laboratories Harrogate (CLEH) standard operating procedures. The analysts were physically located remote from the CLEH facility, but were subject to CLEH management and GLP control systems (including QA inspection). All slides and raw data generated by the remote analysts were returned to CLEH for archiving on completion of analysis.

Analysis of results

Treatment of data

After completion of scoring and decoding of slides, the numbers of binucleate cells with micronuclei (MNBN cells) in each culture were obtained.

The proportions of MNBN cells in each replicate were used to establish acceptable heterogeneity between replicates by means of a binomial dispersion test (Richardson *et al.*, 1989).

The proportion of MNBN cells for each treatment condition were compared with the proportion in negative controls by using Fisher's exact test (Richardson *et al.*, 1989). Probability values of $p \leq 0.05$ were accepted as significant. Additionally, the number of micronuclei per binucleate cell were obtained and recorded.

Acceptance criteria

The assay was considered valid if the following criteria were met:

1. The binomial dispersion test demonstrated acceptable heterogeneity (in terms of MNBN cell frequency) between replicate cultures, particularly where no positive responses were seen.

2. The frequency of MNBN cells in vehicle controls fell within the normal ranges.

3. The positive control chemicals induced statistically significant increases in the proportion of cells with micronuclei. Both replicate cultures at the positive control concentration analysed under each treatment condition demonstrated MNBN cell frequencies that clearly exceeded the current historical vehicle control ranges.

4. A minimum of 50% of cells had gone through at least one cell division (as measured by binucleate + multinucleate cell counts) in negative control cultures at the time of harvest.

Evaluation criteria

For valid data, the test article was considered to induce clastogenic and/or aneugenic events if:

1. A statistically significant increase in the frequency of MNBN cells at one or more concentrations was observed.

2. An incidence of MNBN cells at such a concentration that exceeded the normal range in both replicates was observed.

3. A concentration-related increase in the proportion of MNBN cells was observed.

The test article was considered positive in this assay if all of the above criteria were met.

The test article was considered negative in this assay if none of the above criteria were met.

Results which only partially satisfied the above criteria were dealt with on a case-by-case basis. Evidence of a concentration-related effect was considered useful but not essential in the evaluation of a positive result (Scott *et al.*, 1990).

Computer systems

The major computer systems used on this study were as follows:

Activity	Computer system
Scheduling	CMS (Covance Management System)
Formulations	Pristima
Slide coding and/or data analysis	Vitroabs
Data generation and collation	Vitronuc/ CBPI
Report generation	Microsoft Office/Adobe Acrobat

Version numbers of the systems are held on file at Covance.

RESULTS

Selection of concentrations for micronucleus analysis

The results of the RI determinations from the cytotoxicity Range-Finder Experiment were as follows:

Table 7: Data for 3+21 hour treatments -S-9, Range-Finder- male donors

Dose (µg/mL)	Replicate	Mono	Bi	Multi	Total Number of Cells	RI	Cytotoxicity (%)
Vehicle	A	45	146	9	200	0.82	-
	B	46	141	13	200	0.84	
18.14	A	NS	-	-	-	-	-
30.23	A	NS	-	-	-	-	-
50.39	A	NS	-	-	-	-	-
83.98	A	NS	-	-	-	-	-
140.0	A	NS	-	-	-	-	-
233.3	A	NS	-	-	-	-	-
388.8	A	NS	-	-	-	-	-
648.0	A	NS	-	-	-	-	-
1080	A	19	166	15	200	0.98	0
1800	A	23	164	13	200	0.95	0
3000	A	27	155	18	200	0.96	0
5000	A	34	148	18	200	0.92	0

Table 8: Data for 3+21 hour treatments +S-9, Range-Finder- male donors

Dose (µg/mL)	Replicate	Mono	Bi	Multi	Total Number of Cells	RI	Cytotoxicity (%)
Vehicle	A	60	132	8	200	0.74	-
	B	32	160	8	200	0.88	
18.14	A	NS	-	-	-	-	-
30.23	A	NS	-	-	-	-	-
50.39	A	NS	-	-	-	-	-
83.98	A	NS	-	-	-	-	-
140.0	A	NS	-	-	-	-	-
233.3	A	NS	-	-	-	-	-
388.8	A	NS	-	-	-	-	-
648.0	A	NS	-	-	-	-	-
1080	A	29	160	11	200	0.91	0
1800	A	42	148	10	200	0.84	0
3000	A	30	157	13	200	0.92	0
5000	A	33	155	12	200	0.90	0

NS = Not scored
Mono = mononucleate
Bi = binucleate
Multi = multinucleate
RI = replication index

Table 9: Data for 24+24 hour treatments -S-9, Range-Finder- male donors

Dose (µg/mL)	Replicate	Mono	Bi	Multi	Total Number of Cells	RI	Cytotoxicity (%)
Vehicle	A	33	160	7	200	0.87	-
	B	43	151	6	200	0.82	
18.14	A	NS	-	-	-	-	-
30.23	A	NS	-	-	-	-	-
50.39	A	NS	-	-	-	-	-
83.98	A	NS	-	-	-	-	-
140.0	A	NS	-	-	-	-	-
233.3	A	NS	-	-	-	-	-
388.8	A	NS	-	-	-	-	-
648.0	A	NS	-	-	-	-	-
1080	A	27	166	7	200	0.90	0
1800	A	29	163	8	200	0.90	0
3000	A	42	152	6	200	0.82	3
5000	A	38	157	5	200	0.84	1

NS = Not scored
Mono = mononucleate
Bi = binucleate
Multi = multinucleate
RI = replication index

No marked changes in osmolality or pH were observed at the highest concentration tested (5000 µg/mL) as compared to the concurrent vehicle controls (individual data not reported).

The results of the cytotoxicity Range-Finder Experiment were used to select suitable maximum concentrations for the Micronucleus Experiment.

The results of the RI determinations from the Micronucleus Experiment were as follows:

Table 10: Data for 3+21 hour treatments -S-9, Micronucleus Experiment - male donors

Dose (µg/mL)	Replicate	Mono	Bi	Multi	Total Number of Cells	RI	Cytotoxicity (%)
Vehicle	A	87	373	40	500	0.91	-
	B	93	368	39	500	0.89	
	C	82	382	36	500	0.91	
	D	88	369	43	500	0.91	
1000	A	NS	-	-	-	-	-
	B	NS	-	-	-	-	
2000	A	75	391	34	500	0.92	0
	B	84	390	26	500	0.88	
3000	**A**	**66**	**405**	**29**	**500**	**0.93**	**0 #**
	B	**66**	**399**	**35**	**500**	**0.94**	
4000	**A**	**69**	**394**	**37**	**500**	**0.94**	**0 #**
	B	**59**	**407**	**34**	**500**	**0.95**	
5000	**A**	**59**	**410**	**31**	**500**	**0.94**	**0 #**
	B	**53**	**420**	**27**	**500**	**0.95**	

Table 11: Data for 3+21 hour treatments +S-9, Micronucleus Experiment- male donors

Dose (µg/mL)	Replicate	Mono	Bi	Multi	Total Number of Cells	RI	Cytotoxicity (%)
Vehicle	A	93	381	26	500	0.87	-
	B	94	387	19	500	0.85	
	C	84	374	42	500	0.92	
	D	81	385	34	500	0.91	
1000	A	NS	-	-	-	-	-
	B	NS	-	-	-	-	
2000	A	82	391	27	500	0.89	0
	B	89	381	30	500	0.88	
3000	**A**	**78**	**400**	**22**	**500**	**0.89**	**2 #**
	B	**94**	**385**	**21**	**500**	**0.85**	
4000	**A**	**103**	**382**	**15**	**500**	**0.82**	**5 #**
	B	**88**	**393**	**19**	**500**	**0.86**	
5000	**A**	**100**	**380**	**20**	**500**	**0.84**	**4 #**
	B	**93**	**388**	**19**	**500**	**0.85**	

NS = Not scored
Mono = mononucleate
Bi = binucleate
Multi = multinucleate
RI = replication index

[#] Highlighted concentrations selected for analysis

Table 12: Data for 24+24 hour treatments -S-9, Micronucleus Experiment - male donors

Dose (µg/mL)	Replicate	Mono	Bi	Multi	Total Number of Cells	RI	Cytotoxicity (%)
Vehicle	A	59	361	80	500	1.04	-
	B	59	367	74	500	1.03	
	C	62	368	70	500	1.02	
	D	53	390	57	500	1.01	
500.0	A	NS	NS	-	-	-	-
	B	NS	NS	-	-	-	
1000	A	NS	NS	-	-	-	-
	B	NS	NS	-	-	-	
2000	A	56	375	69	500	1.03	0
	B	56	354	90	500	1.07	
3000	**A**	**58**	**361**	**81**	**500**	**1.05**	**0 #**
	B	**52**	**383**	**65**	**500**	**1.03**	
4000	**A**	**43**	**369**	**88**	**500**	**1.09**	**0 #**
	B	**59**	**380**	**61**	**500**	**1.00**	
5000	**A**	**33**	**407**	**60**	**500**	**1.05**	**0 #**
	B	**35**	**392**	**73**	**500**	**1.08**	

NS = Not scored
Mono = mononucleate
Bi = binucleate
Multi = multinucleate
RI = replication index

[#] Highlighted concentrations selected for analysis

Micronucleus analysis

Raw data

The raw data for the observations on the test article plus positive and negative controls are retained by Covance Laboratories Limited. A summary of the number of cells containing micronuclei is given in Appendix 1.

Validity of study

The data in Appendix 1, Appendix 2, Appendix 3 and Table 10 to Table 12 indicate that:

1) The binomial dispersion test demonstrated acceptable heterogeneity (in terms of MNBN cell frequency) between replicate cultures (Appendix 2)

2) The frequency of MNBN cells in vehicle controls fell within the historical vehicle control (normal) range (Appendix 3)

3) The positive control chemicals induced statistically significant increases in the proportion of MNBN cells with both replicate cultures at each positive control concentration analysed clearly exceeding historical vehicle control ranges (Appendix 1 and Appendix 3).

4) A minimum of 50% of cells had gone through at least one cell division (as measured by binucleate + multinucleate cell counts) in negative control cultures at the time of harvest (Table 10 to Table 12).

Analysis of data

Pulse 3+21 hour treatment of cells with Amyloglycosidase PPY 32789 in the absence and presence of S-9 resulted in frequencies of MNBN cells which were similar to and not significantly ($p \leq 0.05$) higher than those observed in concurrent vehicle controls for the majority of concentrations analysed. The single exception to this was for the lowest concentration analysed (3000 µg/mL) following treatment in the presence of S-9 where a small increase was noted (mean MNBN cell frequency = 0.4% compared to 0.15% in the concurrent vehicle control). However, as the MNBN cell frequency of both replicate cultures at this and all other Amyloglycosidase PPY 32789 treated cultures fell within historical vehicle control (normal) values, this isolated statistical increase was not considered of biological importance (Appendix 1, Appendix 2 and Appendix 3).

Following 24+24 hour –S-9 treatment a further small but statistically significant increase in MNBN cells was noted at the intermediate concentration analysed (4000 µg/mL). However, there were no instances of MNBN cell values exceeding normal values for any test article treated culture (at this or any other concentration), with higher and lower concentrations analysed showing mean MNBN cell values similar to those observed for the vehicle control. As such, this isolated statistical increase was not considered of biological importance (Appendix 1, Appendix 2 and Appendix 3).

CONCLUSION

It is concluded that Amyloglycosidase PPY 32789 did not induce micronuclei in cultured human peripheral blood lymphocytes following treatment in the absence and presence of a rat liver metabolic activation system (S-9). Concentrations were tested up to 5000 μg/mL, a recommended regulatory maximum concentration for *in vitro* cytogenetic assays.

REFERENCES

Aardema M S, Albertini S, Arni P, Henderson L M, Kirsch-Volders M, Mackay J M, Sarriff A M, Stringer D A and Taalman R D F (1998) Aneuploidy: a report of an ECETOC task force. Mutation Research 410, 3-79

Brusick D (1986) Genotoxic effects in cultured mammalian cells produced by low pH treatment conditions and increased ion concentrations. Environ Mutagenesis 8, 879-886

Elhajouji A, Cunha M, Kirsch-Volders M (1998) Spindle poisons can induce polyploidy by mitotic slippage and micronucleate mononucleates in the cytokinesis-block assay. Mutagenesis 13, 193-198

Evans H J and O'Riordan M L (1975) Human lymphocytes for analysis of chromosome aberrations in mutagen tests. Mutation Research 31, 135-148

Fenech M and Morley A A (1985) Measurement of micronuclei in human lymphocytes. Mutation Research 147, 29-36

Fenech M and Morley A A (1986) Cytokinesis-block micronucleus method in human lymphocytes: effect of *in-vivo* ageing and low dose X-irradiation. Mutation Research 161, 193-198

Fenech M (1998) Important variables that influence base-line micronucleus frequency in cytokinesis-blocked lymphocytes – a biomarker for DNA damage in human populations. Mutation Research 404, 155-165

Fenech M, Holland N, Chang W P, Zeiger E, Bonassi S (1999) The HUman MicroNucleus project: an international collaborative study on the use of the micronucleus technique for measuring DNA damage in humans. Mutation Research 428, 271-283

Fenech M, Bonassi S, Turner J, Lando C, Ceppi M, Chang W P, Holland N, Kirsch-Volders M, Zeiger E, Bigatti M P, Bolognesi C, Cao J, De Luca G, Di Giorgio M, Ferguson L R, Fucic A, Lima O G, Hadjidekova VV, Hrelia P, Jaworska A, Joksic G, Krishnaja A P, Lee T K, Martelli A, McKay M J, Migliore L, Mirkova E, Muller W U, Odagiri Y, Orsiere T, Scarfi M R, Silva M J, Sofuni T, Suralles J, Trenta G, Vorobtsova I, Vral A and Zijno A (2003) HUman MicroNucleus project. Intra- and inter-laboratory variation in the scoring of micronuclei and nucleoplasmic bridges in binucleated human lymphocytes. Results of an international slide-scoring exercise by the HUMN project. Mutation Research 534, 45-64

Fenech M (2007) Cytokinesis-block micronucleus cytome assay. Nature Protocols, 2(5), 1084-1104

Galloway S M, Aardema M J, Ishidate M, Ivett J L, Kirkland D J, Morita T, Mosesso P, Sofuni T (1994) Report from working group on *in vitro* tests for chromosomal aberrations. In: Sheila M. Galloway (Ed), Report of the International Workshop on Standardisation of Genotoxicity Test Procedures. Mutation Research 312, 241-261

ICH-S2A (1995) "Specific Aspects of Regulatory Genotoxicity Tests for Pharmaceuticals"

ICH-S2B (1997) "Standard Battery for Genotoxicity Testing of Pharmaceuticals"

Lorge E, Thybaud V, Aardema M J, Oliver J, Wakata A, Lorenzon G and Marzin D (2006) SFTG International collaborative Study on *in vitro* micronucleus test. I. General conditions of the study. Mutation Research 607, 13-36

Migliore L and Nieri M (1991) Evaluation of twelve potential aneuploidogenic chemicals by the *in vitro* human lymphocyte micronucleus assay. Toxicology *In Vitro* 5, 325-336

Miller B, Potter-Locher F, Seelbach A, Stopper H, Utesch D and Madle S (1998) Evaluation of the *in vitro* micronucleus test as an alternative to the *in vitro* chromosomal aberration assay: position of the GUM working group on the *in vitro* micronucleus test. Mutation Research 410, 81-116

Mitelman F (1991) "Catalogue of Chromosome Aberrations in Cancer, 4th ed". New York: Wiley-Liss

OECD (2010) 'Genetic Toxicology: OECD Guideline for the testing of chemicals. Guideline 487: *In vitro* mammalian cell micronucleus test

Preston R J, Au W, Bender M A, Brewen J G, Carrano A V, Heddle J A, McFee A F, Wolff S and Wassom J S (1981) Mammalian *in vivo* and *in vitro* cytogenetic assays. A report of the U.S. EPA's Gene-Tox Program. Mutation Research 87, 143-188

Ramirez M J, Surrallés J, Puerto S, Creus A, Marcos R (1997) Aneugenic activity in human cultured lymphocytes. An overall study with colchicine using the micronucleus assay and fluorescence *in situ* hybridization techniques. Mutagenesis Vol 12, No 6, 405-410

Richardson C, Williams D A, Allen J A, Amphlett G, Chanter D O and Phillips B (1989) Analysis of data from *in vitro* cytogenetic assays. In "Statistical Evaluation of Mutagenicity Test Data", (UKEMS Guidelines Sub-committee Report, Part III), Ed D J Kirkland, Cambridge University Press, pp 141-154

Rosefort C, Fauth E and Zankl H (2004) Micronuclei induced by aneugens and clastogens in mononucleate and binucleate cells using the cytokinesis block assay. Mutagenesis 19, 277-284

Scott D, Dean B J, Danford N D Kirkland D J (1990) Metaphase chromosome aberration assays *in vitro*. Basic Mutagenicity Tests; UKEMS recommended procedures. Kirkland D J (Ed) pp 62-86

Scott D, Galloway S M, Marshall R R, Ishidate M, Brusick D, Ashby J and Myhr B C (1991) Genotoxicity under extreme culture conditions. A report from ICPEMC Task Group 9. Mutation Research 257, 147-204

Thybaud V, Aardema M, Clements J, Dearfield K, Galloway S, Hayashi M, Jacobson-Kram D, Kirkland D, MacGregor J T, Marzin D, Ohyama W, Schuler M, Suzuki H, Zeiger E (2007) Strategy for genotoxicity testing: Hazard identification and risk assessment in relation to *in vitro* testing. Mutation Research 627, 41-58.

APPENDICES

Appendix 1
Binucleate cells with micronuclei

Table 13: Amyloglycosidase PPY 32789, 3+21 hour treatments in the absence of S-9
Micronucleus Experiment - male donors

Treatment (µg/mL)	Replicate	Total BN Cells Scored	Total MNBN Cells Scored	Frequency of MNBN Cells/ Cells Scored (%)	Significance § (% Toxicity)
Vehicle	A	1000	3	0.30	
	B	1000	2	0.20	
	C	1000	5	0.50	
	D	1000	3	0.30	
	Total	4000	13	0.33	-
3000	A	1000	3	0.30	
	B	1000	5	0.50	
	Total	2000	8	0.40	NS (0)
4000	A	1000	4	0.40	
	B	1000	2	0.20	
	Total	2000	6	0.30	NS (0)
5000	A1	1000	10	**1.00 #**	
	A2	1000	3	0.30	
	B1	1000	3	0.30	
	B2	1000	4	0.40	
	Total	4000	20	0.50	NS (0)
MMC, 0.80	A	1000	150	**15.00 #**	
	B	1000	113	**11.30 #**	
	Total	2000	263	13.15	$p \leq 0.001$

MNBN = Micronucleated Binucleate
§ Statistical significance (Appendix 2)
NS = Not significant
# = Numbers highlighted exceed historical negative control range (Appendix 3)

Table 14: Amyloglycosidase PPY 32789, 3+21 hour treatments in the presence of S-9

Micronucleus Experiment - male donors

Treatment (µg/mL)	Replicate	Total BN Cells Scored	Total MNBN Cells Scored	Frequency of MNBN Cells/ Cells Scored (%)	Significance § (% Toxicity)
Vehicle	A	1000	0	0.00	
	B	1000	2	0.20	
	C	1000	2	0.20	
	D	1000	2	0.20	
	Total	4000	6	0.15	-
3000	A	1000	2	0.20	
	B	1000	6	0.60	
	Total	2000	8	0.40	$p \leq 0.05$ (2)
4000	A	1000	4	0.40	
	B	1000	3	0.30	
	Total	2000	7	0.35	NS (5)
5000	A	1000	1	0.10	
	B	1000	3	0.30	
	Total	2000	4	0.20	NS (4)
CPA, 12.5	A	1000	29	**2.90 #**	
	B	1000	41	**4.10 #**	
	Total	2000	70	3.50	$p \leq 0.001$

MNBN = Micronucleated Binucleate
§ Statistical significance (Appendix 2)
NS = Not significant
# = Numbers highlighted exceed historical negative control range (Appendix 3)

Table 15: Amyloglycosidase PPY 32789, 24+24 hour treatments in the absence of S-9

Micronucleus Experiment - male donors

Treatment (µg/mL)	Replicate	Total BN Cells Scored	Total MNBN Cells Scored	Frequency of MNBN Cells/ Cells Scored (%)	Significance § (% Toxicity)
Vehicle	A	1000	2	0.20	
	B	1000	2	0.20	
	C	1000	3	0.30	
	D	1000	3	0.30	
	Total	4000	10	0.25	-
3000	A	1000	2	0.20	
	B	1000	4	0.40	
	Total	2000	6	0.30	NS (0)
4000	A	1000	8	0.80	
	B	1000	4	0.40	
	Total	2000	12	0.60	$p \leq 0.05$ (0)
5000	A	1000	1	0.10	
	B	1000	1	0.10	
	Total	2000	2	0.10	NS (0)
VIN, 0.08	A	1000	74	**7.40 #**	
	B	1000	85	**8.50 #**	
	Total	2000	159	7.95	$p \leq 0.001$

MNBN = Micronucleated Binucleate
§ Statistical significance (Appendix 2)
NS = Not significant
# = Numbers highlighted exceed historical negative control range (Appendix 3)

Appendix 2
Statistical analysis of test article data

Table 16: Amyloglycosidase PPY 32789, 3+21 hour treatments in the absence of S-9
Micronucleus Experiment - male donors

Binomial Dispersion Test $\chi^2 = 9.47$	DF: 8
Significance: NS	

Treatment (µg/mL)	Total BN Cells	BN Cells with micronuclei	Proportion	Fisher's exact test	Significance
Vehicle	4000	13	0.003	-	-
3000	2000	8	0.004	0.319	NS
4000	2000	6	0.003	0.553	NS
5000	2000	20	0.005	0.114	NS
MMC, 0.80	4000	263	0.132	0.000	$p \leq 0.001$

Table 17: Amyloglycosidase PPY 32789, 3+21 hour treatments in the presence of S-9
Micronucleus Experiment - male donors

Binomial Dispersion Test $\chi^2 = 5.16$	DF:6
Significance: NS	

Treatment (µg/mL)	Total BN Cells	BN Cells with micronuclei	Proportion	Fisher's exact test	Significance
Vehicle	4000	6	0.002	-	-
3000	2000	8	0.004	0.037	$p \leq 0.05$
4000	2000	7	0.004	0.069	NS
5000	2000	4	0.002	0.327	NS
CPA, 12.5	2000	70	0.035	0.000	$p \leq 0.001$

NS = Not significant DF = degrees of freedom BN = binucleate

Table 18: Amyloglycosidase PPY 32789, 24+24 hour treatment in the absence of S-9

Micronucleus Experiment - male donors

Binomial Dispersion Test $\chi^2 = 2.41$	DF: 6
Significance: NS	

Treatment (µg/mL)	Total BN Cells	BN Cells with micronuclei	Proportion	Fisher's exact test	Significance
Vehicle	4000	10	0.003	-	-
3000	2000	6	0.003	0.358	NS
4000	2000	12	0.006	0.022	$p \leq 0.05$
5000	2000	2	0.001	0.883	NS
VIN, 0.08	2000	159	0.080	0.000	$p \leq 0.001$

NS = Not significant DF = degrees of freedom BN = binucleate

Appendix 3
Historical vehicle control ranges for the human peripheral blood lymphocyte micronucleus assay

Table 19: Historical vehicle control range - 3+21 hour

		Micronucleated binucleates observed in 1000 binucleates scored Male donors	Frequency of MNBN cells/cells scored (%) Male donors
-S9	Number of studies	5	5
	Number of cultures	42	42
	Median	5	0.45
	Mean	5	0.45
	SD	1.8	0.176
	Observed range	1 - 9	0.10 - 0.90
	95% reference range	2 - 8	0.2-0.8
	99% reference range	NR	NR
+S9	Number of studies	5	5
	Number of cultures	42	42
	Median	4	0.40
	Mean	5	0.48
	SD	2.8	0.282
	Observed range	1 - 11	0.10 - 1.1
	95% reference range	1 - 11	0.1 – 1.1
	99% reference range	NR	NR

Reference ranges are calculated from percentiles of the observed distributions.

Data generated from QA audited data.

Calculated in May 2011 by CLEH Statistics, for studies started between January 2011 and April 2011.

Table 20: Historical vehicle control range - 24+0 hour

		Micronucleated binucleates observed in 1000 binucleates scored Male donors	Frequency of MNBN cells/cells scored (%) Male donors
-S9	Number of studies	7	7
	Number of cultures	53	53
	Median	5	0.50
	Mean	5	0.48
	SD	2.6	0.259
	Observed range	0 - 13	0.00 - 1.30
	95% reference range	0 - 11	0 – 1.1
	99% reference range	NR	NR

Reference ranges are calculated from percentiles of the observed distributions.

Data generated from QA audited data.

Calculated in May 2011 by CLEH Statistics, for studies started between January 2011 and April 2011.

Appendix 4
Quality control statements for S-9

MOLTOX POST MITOCHONDRIAL SUPERNATANT (S-9)
QUALITY CONTROL & PRODUCTION CERTIFICATE

LOT NO.: 2845
PART NO.: 11-101
VOLUME: 5mL
SPECIES: Rat
STRAIN: Sprague Dawley
SEX: Male
TISSUE: Liver
PREPARATION DATE: October 11, 2011
EXPIRATION DATE: October 11, 2013
BUFFER: 0.154 M KCl
INDUCING AGENT(s): Aroclor 1254 (Monsanto KL615), 500 mg/kg i.p.
REFERENCE: Maron, D & Ames, B, *Mutat Res* 113:173, 1983
STORAGE: At or below -70°C

BIOCHEMISTRY:

- PROTEIN

44.3 mg/ml — Assayed according to the method of Lowry et al., *JBC* 193:265, 1951 using bovine serum albumin as the standard.

- ALKOXYRESORUFIN-0-DEALKYLASE ACTIVITIES

Activity	P450	Fold - Induction
EROD	1A1, 1A2	254.9
PROD	2B1, 2B2	65.3
BROD	2B1, 2B2	92.7
MROD	1A1, 1A2	109.4

Assays for ethoxyresorufin-0-deethylase (EROD), pentoxy-, benzyl- and methoxyresorufin-0-dealkylases (PROD, BROD, & MROD) were conducted using a modification of the methods of Burke, et al., *Biochem Pharm* 34:3337, 1985. Fold-inductions were calculated as the ratio of the sample vs. uninduced specific activities (SA's). Control SA's (pmoles/min/mg protein) were 40.0, 13.4, 55.6, & 29.0 for EROD, PROD, BROD and MROD, respectively.

BIOASSAY:

- TEST FOR THE PRESENCE OF ADVENTITIOUS AGENTS

Samples of S-9 were assayed for the presence of contaminating microflora by plating 1.0 ml volumes on Nutrient Agar and Minimal Glucose (Vogel-Bonner E, supplemented with 0.05 mM L-histidine and D-biotin) media. Triplicate plates were read after 40 - 48 h incubation at 35 ± 2°C. The tested samples met acceptance criteria.

- PROMUTAGEN ACTIVATION

No. His+ Revertants	
TA98	TA1535
167.6	1074

The ability of the sample to activate ethidium (EtBr) EtBr/CPA/and cyclophosphamide (CPA) to intermediates mutagenic to TA98 and TA1535, respectively, was determined according to Lesca, et al., *Mutation Res* 129:299, 1984. Data were expressed as revertants per µg EtBr or per mg CPA.

Dilutions of the sample S9, ranging from 0.2 - 10% in S9 mix, were tested for their ability to activate benzo(a)pyrene (BP) and 2-aminoanthracene (2-AA) to intermediates mutagenic to TA100. Assays were conducted as described by Maron & Ames, (*Mutat Res* 113:173, 1983).

µl S9 per plate/number *his* revertants per plate

Promutagen	0	1	5	10	20	50
BP (5 µg)	106	379	605	728	858	545
2-AA (2.5 µg)	118	1767	2183	2077	1301	459

MOLECULAR TOXICOLOGY, INC.
157 Industrial Park Dr.
Boone, NC 28607
(828) 264-9099
www.moltox.com

9/11

11-21-11 fn 11-4-11

MOLTOX POST MITOCHONDRIAL SUPERNATANT (S-9) QUALITY CONTROL & PRODUCTION CERTIFICATE

LOT NO.: 2854
PART NO.: 11-101
VOLUME: 5mL
SPECIES: Rat
STRAIN: Sprague Dawley
SEX: Male
TISSUE: Liver
PREPARATION DATE: November 02, 2011
EXPIRATION DATE: November 02, 2013
BUFFER: 0.154 M KCl
INDUCING AGENT(s): Aroclor 1254 (Monsanto KL615), 500 mg/kg i.p.
REFERENCE: Maron, D & Ames, B, *Mutat Res* **113**:173, 1983
STORAGE: At or below -70°C

BIOCHEMISTRY:

- PROTEIN

44.3 mg/ml

Assayed according to the method of Lowry et al., *JBC* **193**:265, 1951 using bovine serum albumin as the standard.

- ALKOXYRESORUFIN-0-DEALKYLASE ACTIVITIES

Activity	P450	Fold - Induction
EROD	1A1, 1A2	341.7
PROD	2B1, 2B2	59.6
BROD	2B1, 2B2	131.0
MROD	1A1, 1A2	104.5

Assays for ethoxyresorufin-0-deethylase (EROD), pentoxy-, benzyl- and methoxyresorufin-0-dealkylases (PROD, BROD, & MROD) were conducted using a modification of the methods of Burke, et al., *Biochem Pharm* **34**:3337, 1985. Fold-inductions were calculated as the ratio of the sample vs. uninduced specific activities (SA's). Control SA's (pmoles/min/mg protein) were 35.7, 17.4, 43.4, & 30.3 for EROD, PROD, BROD and MROD, respectively.

BIOASSAY:

- TEST FOR THE PRESENCE OF ADVENTITIOUS AGENTS

Samples of S-9 were assayed for the presence of contaminating microflora by plating 1.0 ml volumes on Nutrient Agar and Minimal Glucose (Vogel-Bonner E, supplemented with 0.05 mM L-histidine and D-biotin) media. Triplicate plates were read after 40 - 48 h incubation at 35 ± 2°C. The tested samples met acceptance criteria.

- PROMUTAGEN ACTIVATION

No. His+ Revertants	
TA98	TA1535
142.8	964

The ability of the sample to activate ethidium (EtBr) EtBr/CPA/and cyclophosphamide (CPA) to intermediates mutagenic to TA98 and TA1535, respectively, was determined according to Lesca, et al., *Mutation Res* **129**:299, 1984. Data were expressed as revertants per µg EtBr or per mg CPA.

Dilutions of the sample S9, ranging from 0.2 - 10% in S9 mix, were tested for their ability to activate benzo(a)pyrene (BP) and 2-aminoanthracene (2-AA) to intermediates mutagenic to TA100. Assays were conducted as described by Maron & Ames, (*Mutat Res* **113**:173, 1983).

µl S9 per plate/number *his* revertants per plate

Promutagen	0	1	5	10	20	50
BP (5 µg)	178	360	451	516	860	841
2-AA (2.5 µg)	147	1077	1959	2038	1616	754

MOLECULAR TOXICOLOGY, INC.
157 Industrial Park Dr.
Boone, NC 28607
(828) 264-9099
www.moltox.com

Approved: 11/04/11

11-21-11 for 11-4-11

Appendix 5
Documentation of Test Material

novozymes
Rethink Tomorrow

Toxicology

Date: 05. October, 2011
Project no.: OPT 05330
Luna: 2011-16779-01
Ref.: KM

Documentation of Test Material

Product:	**TOX BATCH**
Batch:	**PPY 32789**
Type of enzyme:	**Amyloglucosidase**
Host organism:	***Aspergillus niger***
Physical form / Colour:	**Brownish liquid at room temperature**
E.C.:	**3.2.1.3**

Activity:	**392 AGU/g**
Water (KF):	**89.8 % w/w**
Dry matter:	**10.2 % w/w**
Ash (600°C):	**0.9 % w/w**
Total Organic Solids (TOS):	**9.3 % w/w**
Specific gravity (g/ml):	**1.042 g/ml**
pH:	**4.8**
Total viable counts/g:	**<100**

Study Director

T-AMG, PPY 24900

TOXICITY STUDY BY

ORAL ADMINISTRATION TO CD RATS FOR 13 WEEKS

Sponsor

Novozymes A/S
Krogshoejvej 36
DK-2880 Bagsvaerd
DENMARK

Research Laboratory

Huntingdon Life Sciences Ltd.,
Woolley Road,
Alconbury,
Huntingdon,
Cambridgeshire,
PE28 4HS,
ENGLAND.

Final: 5 October 2006

CONTENTS

CONTENTS - continued

Page

COMPLIANCE WITH GOOD LABORATORY PRACTICE STANDARDS

T-AMG, PPY 24900

TOXICITY STUDY BY

ORAL ADMINISTRATION TO CD RATS FOR 13 WEEKS

The aspects of the study that were performed at Huntingdon Life Sciences were conducted in compliance with the following Good Laboratory Practice standards and I consider the data generated to be valid:

> The UK Good Laboratory Practice Regulations (Statutory Instrument 1999 No. 3106, as amended by Statutory Instrument 2004 No. 994).
>
> OECD Principles of Good Laboratory Practice (as revised in 1997), ENV/MC/CHEM(98)17.
>
> EC Commission Directive 2004/10/EC of 11 February 2004 (Official Journal No. L 50/44).

The investigations performed at the Enzyme Analytical Laboratory, Novozymes A/S, were conducted in compliance with the OECD Principle of Good Laboratory Practice, ENV/MC/CHEM (98) 17.

These principles of Good Laboratory Practice are accepted by the regulatory authorities of the United States of America and Japan on the basis of intergovernmental agreements.

..... 5 October 2006.....

Date

Study Director,
Huntingdon Life Sciences Ltd.

QUALITY ASSURANCE STATEMENT

T-AMG, PPY 24900

TOXICITY STUDY BY

ORAL ADMINISTRATION TO CD RATS FOR 13 WEEKS

The following inspections and audits have been carried out in relation to this study:

Study Phase	Date(s) of Inspection	Date of Reporting to Study Director and Management
Protocol Audit	24 November 2005	24 November 2005
Study Based Inspections		
Formulation Procedures	25-28 November 2005	28 November 2005
Dosing, Post Dose Observations and Study Preparation	28 November 2005	28 November 2005
Clinical Signs	9 December 2005	9 December 2005
Necropsy	27 February 2006	27 February 2006
Report Audit	16-21 June 2006	21 June 2006

In addition, process based inspections were conducted of other routine and repetitive procedures employed on this type of study at or about the time this study was in progress. Similarly an inspection of the facility where this study was conducted was carried out on an annual basis. These inspections were reported to Company Management.

Details of Quality Assurance inspections and audits relating to the formulation chemistry phase are indicated in the Test Site QA statement within the Principal Investigator's report included as Annex 1.

..... 5 OCTOBER 2006

Date

Department of Quality Assurance,
Huntingdon Life Sciences Ltd.

CONTRIBUTING SCIENTISTS

T-AMG, PPY 24900

TOXICITY STUDY BY

ORAL ADMINISTRATION TO CD RATS FOR 13 WEEKS

STUDY MANAGEMENT

N. Hughes, HNC
Study Director

A. Broadmeadow, B.Tech. (Hons.), Dip.R.C.Path(Tox)., Eurotox Registered Toxicologist
Senior Consultant Toxicologist

FORMULATION ANALYSIS

M. C. Tonsgaard
Principal Investigator
Novozymes A/S

SENSORY REACTIVITY, GRIP STRENGTH AND MOTOR ACTIVITY

M. J. Collier, B.Sc.
Behavioural Scientist

OPHTHALMIC EXAMINATION

P. Lee, B.V.Sc., M.R.C.V.S., D.V.S.M.
Veterinary Officer

CLINICAL PATHOLOGY

P. Travis B.Sc. (Hons.), M.Sc.
Head of Central Laboratory Services

PATHOLOGY

Woo-Chan Son, D.V.M., M.Sc. (Korea), Ph.D.
Pathologist

STATISTICS

G. F. Healey, B.Sc. (Hons.), M.Sc., A.R.C.S.
Head of Statistics

SUMMARY

The systemic toxic potential of T-AMG, PPY 24900 (an enzyme intended for use in the food industry) to Crl:CD® (SD) BR rats by oral administration was assessed over a period of 13 weeks. Three groups, each comprising ten males and ten females, received T AMG, PPY 24900 at doses of 1.0, 3.3 or 10.0 mL/kg/day (equivalent to 0.15, 0.48 or 1.47 gTOS/kg/day or 553, 1824 or 5528 AGU/kg/day). A similarly constituted Control group received the vehicle (purified water obtained by reverse osmosis) at the same volume-dosage.

During the study, clinical condition, detailed physical and arena observations, sensory reactivity, grip strength, motor activity, bodyweight, food and water consumption, ophthalmic examination, haematology, blood chemistry, organ weight, macropathology and histopathology investigations were undertaken.

Results

There were no signs related to treatment seen at the routine weekly physical examinations. There were no toxicologically significant findings at the behavioural investigations.

There were no treatment-related deaths during the course of the treatment period. One male given 3.3 mL/kg/day died through a cause that was not related to treatment.

There was no toxicologically significant effect upon bodyweight gain, food and water consumption and food conversion efficiency.

There were no treatment-related ophthalmic findings.

Haematological and blood biochemical investigations in Week 13 did not indicate any changes of toxicological significance.

Organ weights and macroscopic pathology were not affected by treatment.

The histopathological examination indicated an increased incidence or increased severity of cortical vacuolation in the adrenal glands of males given 3.3 or 10.0 mL/kg/day, but this was considered a slight exacerbation of a normal background finding in young CD rats which was not considered toxicologically significant.

Conclusion

It is concluded that oral administration of T-AMG, PPY 24900 (a food enzyme) to CD rats at doses up to 10 mL/kg/day for 13 weeks was well tolerated and did not cause any toxicologically significant change. In males receiving 3.3 or 10 mL/kg/day there was a small increased in incidence and severity of adrenal cortical vacuolation but this was not considered toxicologically significance. Consequently, the no-observed-adverse-effect level (NOAEL) in this study was considered to be 10 mL/kg/day, which is equivalent to 5528 AGU/kg/day or 1.47 gTOS/kg/day.

INTRODUCTION

Objective

The objective of this study was to assess the systemic toxic potential of T-AMG, PPY 24900, a food enzyme, when administered orally by gavage to rats for 13 weeks.

Regulatory compliance

The study was designed to meet the requirements of the following guideline:

> Organisation for Economic Co-operation and Development, Testing of Chemicals Guideline No. 408 (revised 1997).

The study was conducted in accordance with the requirements of current, internationally recognised Good Laboratory Practice Standards, and the applicable sections of the United Kingdom Animals (Scientific Procedures) Act 1986.

Test system

The rat was chosen as the test species because of its acceptance as a predictor of toxic change in man and the requirement for a rodent species by regulatory agencies. The Crl:CD® (SD) BR strain was used because of the historical control data available in this laboratory.

Route of administration

The oral route of administration was chosen to simulate the conditions of potential human exposure.

Treatment groups and doses

The doses used in this study (0, 1.0, 3.3 and 10.0 mL/kg/day) were selected in conjunction with the Sponsor on the basis of results from studies performed on other similar enzyme preparations. The highest dose (10 mL/kg/day) was the maximum practical dose and represents administration of the enzyme, as received, at a volume-dose of 10 mL/kg bodyweight, which is the maximum practical volume-dose for repeat dose oral administration. The lower doses were selected using an approximate ratio of 3.3 between doses.

Study location

The test system was maintained at the following laboratory:

Huntingdon Life Sciences Ltd.,
Eye Research Centre,
Eye,
Suffolk,
IP23 7PX,
England.

The analyses described in the blood chemistry and haematology (with the exception of coagulation), sections of this report were performed by:

Huntingdon Life Sciences Ltd.,
Huntingdon Research Centre,
Woolley Road,
Alconbury,
Huntingdon,
Cambridgeshire,
PE28 4HS,
England.

The analyses of formulations were performed by:

Novozymes A/S
Enzyme Analytical Laboratory, 6E1. 16
Krogshojvej 36
DK-2880 Bagsvaerd
Denmark.

EXPERIMENTAL PROCEDURE

STUDY SCHEDULE AND STRUCTURE

Duration of treatment

The test substance, T-AMG, PPY 24900, was administered over a period of 13 consecutive weeks. The necropsy procedures were completed in two days, during which time treatment continued, and serial observations were recorded at appropriate intervals. The duration of treatment is reported as 13 weeks.

Time schedule

Study initiation: (Protocol signed by Study Director)	16 November 2005
Experimental start date: (animal arrival)	16 November 2005
Treatment commenced:	28 November 2005
Necropsy completed:	28 February 2006
Experimental completion date: (pathology	26 May 2006
Study completion:	5 October 2006

Identity of treatment groups

The study consisted of one Control and three treated groups of rats, identified as follows:

Group	Treatment	Dose# (mL/kg/day)	Dose (g TOS/kg/day)	Dose (AGU/kg/day)
1	Control	0	0	0
2	T-AMG, PPY 24900	1.0	0.15	553
3	T-AMG, PPY 24900	3.3	0.48	1824
4	T-AMG, PPY 24900	10.0	1.47	5528

Using a volume-dose of 10 mL/kg bodyweight

Group	Treatment	No. of animals		Animal numbers		Cage numbers	
		Male	Female	Male	Female	Male	Female
1	Control	10	10	11-20	51-60	3-4	11-12
2	T-AMG, PPY 24900	10	10	31-40	41-50	7-8	9-10
3	T-AMG, PPY 24900	10	10	1-10	61-70	1-2	13-14
4	T-AMG, PPY 24900	10	10	21-30	71-80	5-6	15-16

Some serial observations needed to be performed without the knowledge of the treatment group, therefore the animal numbering system was such that it was not easy to identify a treatment group from the animal number.

TEST SUBSTANCE AND FORMULATION

Test substance

Information supplied by the Sponsor regarding the test substance is contained in the test substance data sheet, which is retained in study records, and the Certificate of Analysis, which is presented in Annex 2.

The following information is given in summary:

Identification:	T-AMG, PPY 24900
Host organism:	*Aspergillus niger*
Action:	An enzyme preparation for the food industry
Description:	Brownish liquid at room temperature
Storage conditions:	Deep-frozen (approximately -20°C). Thawed and refrozen on no more than one occasion.
Supplier:	Sponsor
Batch number:	PPY 24900
Date of receipt:	21 November 2005
Quantity received:	56 plastic bottles each containing approximately 250 mL
Expiry date:	10 May 2015
Enzyme activity:	524 AGU/g
Water (KF):	85.6% w/w
Dry matter:	14.4% w/w
Ash content (600°C)	0.5% w/w
Total Organic Solid (TOS) content:	13.9% w/w
Specific gravity:	1.055 g/mL
pH:	6.1
Total viable counts/g:	<200

The Sponsor was responsible for the characterisation of the test substance and the documentation of the methods of synthesis, fabrication or derivation and stability.

Following the commencement of treatment, a representative sample of the test substance was taken. This sample was placed in a well-closed glass container and stored in the archives at approximately -20°C.

Vehicle

Purified water obtained by reverse osmosis

Formulation

Prior to the commencement of treatment, and at appropriate intervals during the study, containers of T-AMG, PPY 24900 were thawed in a refrigerator (at approximately 4ºC) and divided to provide aliquots of sufficient size for the required formulation. These aliquots were then re-frozen pending use. The test material was, therefore, thawed and re-frozen on only one occasion.

At appropriate intervals before and during the study (twice each week as a minimum schedule) the required aliquots of T-AMG, PPY 24900 were removed from the freezer and allowed to thaw overnight in a refrigerator (approximately 4ºC). Formulations were subsequently prepared as follows:

The test enzyme was prepared for administration as a series of graded concentrations in purified water (obtained by reverse osmosis) to provide the required dosages at a constant volume-dosage.

Group	Treatment	Dose (mL/kg/day)	Concentration % material as supplied	Volume dosage (mL/kg)
1	Control	0	0	10
2	T-AMG, PPY 24900	1.0	10	10
3	T-AMG, PPY 24900	3.3	33	10
4	T-AMG, PPY 24900	10.0	100	10

Each formulation for the intermediate and low dosage groups was prepared by gentle magnetic stirring of the total required quantity of T-AMG, PPY 24900 and vehicle. The high dosage group received the test material as supplied. Following preparation, formulations were refrigerated (approximately 4ºC) until required for use. Formulations were administered within the specified stability period (i.e. 96 hours).

Quality control of dose form

Detailed records of compound usage were maintained. The amount of test substance necessary to prepare the formulations and the amount actually used were determined on each occasion. The difference between these amounts was checked before the formulations were dispensed.

The suitability of the proposed mixing procedure was determined as part of a separate study performed by the Sponsor. Assessment of homogeneity was not relevant because the liquid enzyme is completely miscible with water. The Sponsor was responsible for the stability of the test enzyme in the liquid matrix. Samples of each formulation prepared for

administration in Weeks 1, 6 and 13 of treatment were analysed with regard to the enzyme activity of the test substance, by the Principal Investigator, after completion of the study. Results of these analyses are presented in Annex 1.

ANIMAL MANAGEMENT

Animal supply, acclimatisation and allocation

A total of 45 male and 45 female Crl:CD® (SD) BR rats were received from Charles River (UK) Ltd. The rats were ordered at 33 to 37 days of age and within a weight range of 15 g for each sex.

On arrival, the animals were removed from the transit boxes and non-selectively allocated to study cages. Using the sequence of cages in the battery, one animal at a time was placed in each cage with the procedure being repeated until each cage held the appropriate number of animals. Each sex was allocated separately.

The cages constituting each group were blocked together by sex and the groups were dispersed in batteries so that possible environmental influences arising from their spatial distribution were equilibrated, as far as was practicable. Additionally, batteries of cages were rotated around the room at weekly intervals to further minimise possible spatial variations.

Each animal was assigned a number and identified uniquely within the study by a tail tattoo. Each cage label was colour-coded according to group and was numbered uniquely with cage and study number, as well as the identity of the occupants.

Before the start of treatment, three males with bodyweights at the extreme of the weight range were replaced with spare animals of suitable weight from the same batch. On Day 1 (before dosing) the bodyweights of the animals were checked to ensure that variations in bodyweight did not exceed ± 20% of the mean for each sex. No animals were outside the specified weight range and therefore no replacements were made.

The animals were allowed to acclimatise to the conditions described below for 12 days before treatment commenced. For those animals selected for this study, their age at the start of treatment was 45 to 49 days and their bodyweights were in the range of 230 to 293 g for males and 150 to 191 g for females.

The spare animals were removed from the study room after treatment commenced.

Animal housing, diet and water supply

Animals were housed inside a barriered rodent facility (Building 30, Room 3019). The facility was designed and operated to minimise the entry of external biological and chemical agents and to minimise the transference of such agents between rooms. Before the study the room was cleaned and disinfected with a bactericide.

Each animal room was kept at positive pressure with respect to the outside by its own supply of filtered fresh air, which was passed to atmosphere and not re-circulated. The temperature and relative humidity controls were maintained within the range of 19 to 23°C and 40 to 70% respectively. Artificial lighting was controlled to give a cycle of 12 hours continuous light and 12 hours continuous dark per 24 hours.

Periodic checks were made on the number of air changes in the animal rooms. Temperature and humidity were monitored daily. There were no significant deviations from target values (see deviations from protocol section).

Alarms were activated if there was any failure of the ventilation system, or temperature limits were exceeded. A stand-by electricity supply was available to be automatically brought into operation should the public supply fail.

The animals were housed five of one sex per cage, unless this number was reduced by mortality. The cages were made of a polycarbonate body with a stainless steel mesh lid. The cages had wood shavings as bedding (Lignocel type 3/4 wood flakes). In addition, each cage was provided with Aspen chew blocks supplied by Datesand which were replaced when necessary. Prior to clinical pathology investigations blocks were removed for the same period as the food. Cages, food hoppers and water bottles were changed at appropriate intervals.

The animals were allowed free access to a standard rodent diet (Rat and Mouse No. 1 Maintenance Diet from Special Diets Services Ltd., Witham, Essex, England), except overnight before routine blood sampling. This diet contained no added antibiotic or other chemotherapeutic or prophylactic agent.

Potable water taken from the public supply was freely available via polycarbonate bottles fitted with sipper tubes.

Each batch of diet was analysed routinely by the supplier for various nutritional components and chemical and microbiological contaminants. Supplier's analytical certificates were scrutinised and approved before any batch of diet was released for use. Each batch of wood flakes was analysed routinely by the supplier for various chemical and microbiological contaminants. The quality of the water supply is governed by regulations published by the Department for Environment, Food and Rural Affairs. Certificates of analysis were received routinely from the water supplier. Certificates of analysis were received routinely from the supplier of the aspen chew blocks. Since the results of these various analyses did not provide evidence of contamination that might have prejudiced the study, they are not presented.

No other specific contaminants that were likely to have been present in the wood shavings, diet or water were analysed, as none that may have interfered with or prejudiced the outcome of the study was known.

Administration

Animals received the test substance or vehicle control formulations orally at a volume-dose of 10 mL/kg bodyweight, using a suitably graduated syringe and a rubber catheter inserted via the mouth into the stomach.

All animals were dosed in sequence of cage-number within each group, once each day at approximately the same time each day, seven days per week. The volume administered to each animal was calculated from the most recently recorded bodyweight.

A daily record of the weight of each formulation dispensed and the amount remaining after dosing was made. The balance of these two weights was compared with the predicted usage as a check that the doses had been administered correctly. No significant discrepancy was found.

Formulations were stirred gently, in order not to damage the enzyme protein, using a magnetic stirrer before and throughout the dosing procedure.

SERIAL OBSERVATIONS

Dated and signed records of all activities relating to the day by day running and maintenance of the study within the animal unit as well as to the group observations and examinations outlined in this experimental procedure were recorded in the Study Day Book. In addition, observations relating to individual animals made throughout the day were recorded.

All observations described below were performed in cage number sequence except where otherwise indicated.

Clinical observations

Animals were inspected visually at least twice daily for evidence of ill-health or reaction to treatment. Cages were inspected daily for evidence of ill-health amongst the occupants. Any deviation from normal was recorded at the time in respect of nature and severity, date and time of onset, duration and progress of the observed condition, as appropriate.

Daily during the first week of treatment, twice weekly during Weeks 2 to 4 (middle and end of each week) and weekly thereafter, detailed observations were recorded at the following times in relation to dose administration:

Immediately before dosing
Immediately after dosing on return of the animal to its cage
On completion of dosing of each group
Between one and two hours after completion of dosing of all groups
As late as possible in the working day.

Detailed physical examination and arena observations

Before treatment commenced and during each week of treatment, detailed physical examination and arena observations were performed on each animal. On each occasion, the examinations were performed at approximately the same time of day (before dosing during the treatment period) by an observer unaware of the experimental group to which the animal belonged.

After removal from the home cage, animals were assessed for physical condition and behaviour during handling and after being placed in a standard arena. Any deviation from normal was recorded with respect to the nature and, where appropriate, degree of severity. Particular attention was paid to possible signs of neurotoxicity.

Findings were either reported as "present" or assigned a severity grade - slight, moderate or marked.

Sensory reactivity and grip strength

During Week 12 of treatment (before dosing), sensory reactivity and grip strength assessment was performed by an observer unaware of the experimental group to which each animal belonged. Animals were not necessarily all tested on the same day, but the number of animals was balanced across the groups on each day of testing. These observations were performed before any laboratory investigations.

The following measurements, reflexes and responses were recorded:

Approach response - A blunt probe was brought towards the animal's head until it was close to the animal's nose (but not touching the vibrissae). The animals reaction was recorded as:

1 - no reaction or ignores probe;
2 - normal awareness and reaction (approaches and/or sniffs probe);
3 - abnormally fearful or aggressive reaction.

Grip strength - Forelimb and hindlimb grip strength was measured using Mecmesin Portable Force Indicators. Three trials were performed.

Auditory startle reflex - The animal's response to a sudden loud noise was assessed. The animal was stationary and the source of sound was not visible. The response was scored as:

1 - no response;
2 - weak response (ear twitch only);
3 - normal response (obvious flinch or startle);
4 - exaggerated response (all feet off the floor).

Tail pinch response - The animal's tail was pinched sharply with forceps approximately one third from the tip. The response was graded as:

1 - no response;
2 - weak response (e.g. turns round slowly or weak vocalisation without moving away);
3 - normal response (e.g. jumps forward or turns around sharply, usually with vocalisation);
4 - exaggerated response (e.g. excessive vocalisation, body movement or aggression).

Touch response - The nape of the animal's neck was stroked gently with a blunt probe and the reaction recorded as:

1 - no reaction or ignores probe;
2 - normal awareness and reaction;
3 - abnormally fearful or aggressive reaction.

Motor activity

During Week 12 of treatment (before dosing), the motor activity of each animal was measured using a Rodent Activity Monitoring System, with hardware supplied by Pearson Technical Services (Saxmundham, Suffolk, England) and software developed and maintained by Huntingdon Life Sciences.

Animals were tested individually in clear polycarbonate cages and motor activity was measured by counting infra-red beam breaks over ten 6-minute intervals (one hour total). Ten beams were set at two height levels (five low and five high) to detect cage floor and rearing activity. All animals were not necessarily tested on the same day, but the number of animals was balanced across the groups on each day of testing. These observations were performed before any laboratory investigations.

Mortality

One male (Animal No. 10) was killed for welfare reasons during Week 3 of treatment; a complete necropsy was performed.

Bodyweight

The weight of each rat was recorded on the day that treatment commenced (Week 0), weekly throughout the treatment period, and before necropsy.

More frequent weighings were instituted, when appropriate, for animals displaying ill-health, so that the progress of the observed condition could be monitored. These data are not reported here.

Food consumption

The weight of food supplied to each cage, that remaining and an estimate of any spilled was recorded for each week throughout the treatment period. From these records the mean weekly consumption per animal (g/rat/week) was calculated for each cage.

Water consumption

Weekly water consumption (over a 3-day period in each week) was recorded by weight for each cage of animals, using water bottles fitted with sipper tubes.

Ophthalmic examination

Before treatment commenced, the eyes of all animals allocated to the study (including spare animals) were examined by means of a binocular indirect ophthalmoscope. During Week 13 of treatment the eyes of all animals of Groups 1 (Control) and 4 (10.0 mL/kg/day) were similarly examined.

Prior to each examination, the pupils of each animal were dilated using 0.5% tropicamide ophthalmic solution (Mydriacyl, Alcon Laboratories Ltd.). The adnexae, conjunctiva, cornea, sclera, anterior chamber, iris (pupil dilated), lens, vitreous and fundus were examined.

As no treatment-related changes were observed, the examination was not extended to animals of Groups 2 or 3 (1.0 or 3.3 mL/kg/day).

Haematology, peripheral blood

During Week 13 of treatment (before dosing), blood samples were obtained from all animals after overnight withdrawal of food. Animals were held under light general anaesthesia induced by isoflurane and blood samples were withdrawn from the sublingual vein.

Blood samples (nominally 0.5 mL) were collected into tubes containing EDTA as anticoagulant and examined for the following characteristics:

The following were measured using a Bayer Advia 120 haematology analyser:

Haematocrit (Hct)
Haemoglobin (Hb)
Erythrocyte count (RBC)
Mean cell haemoglobin (MCH)
Mean cell haemoglobin concentration (MCHC)
Mean cell volume (MCV)
Total white cell count (WBC)
Differential WBC count
- Neutrophils (N)
- Lymphocytes (L)
- Eosinophils (E)
- Basophils (B)
- Monocytes (M)
- Large unstained cells (LUC)

Platelet count (Plt)

Morphology flags were generated by the Advia 120 analyser. The most common morphological changes, anisocytosis, micro/macrocytosis and hypo/hyperchromasia were recorded as follows:

-	no abnormalities detected
+	slight
++	moderate
+++	marked

Blood film (prepared for all samples) - Romanowsky stain, examined for abnormalities by light microscopy, in the case of flags from the Advia 120 analyser. Confirmation or a written description from the blood film was made where appropriate.

Additional blood samples (nominally 0.5 mL) were taken into tubes containing citrate anticoagulant and examined in respect of:

Prothrombin time (PT) - using an ACL 3000 Analyser and IL PT-Fibrinogen reagent

Activated partial thromboplastin time (APTT) - using an ACL 3000 Analyser and IL APTT reagent.

Blood chemistry

At the same time and using the same animals as for peripheral haematology, further blood samples (nominally 0.7 mL) were collected into tubes containing lithium heparin as anticoagulant. All tubes were mechanically agitated for at least two minutes and the sample subsequently centrifuged at 3000 rpm for 10 minutes in order to separate the plasma. After separation, the plasma was examined in respect of:

Using a Hitachi 917 Clinical Chemistry Analyser:

Alkaline phosphatase (ALP)
Alanine aminotransferase (ALT)
Aspartate aminotransferase (AST)
Gamma-glutamyl transpeptidase (gGT)
Total bilirubin (Bili)
Urea
Creatinine (Creat)
Glucose (Gluc)
Total cholesterol (Chol)
Triglycerides (Trig)
Sodium (Na)
Potassium (K)
Chloride (Cl)
Calcium (Ca)
Inorganic phosphorus (Phos)
Total protein (Total Prot)
Albumin (Alb)

Albumin/globulin ratio (A/G Ratio) was calculated from total protein concentration and analysed albumin concentration.

Serum sample collection

At necropsy, without overnight deprivation of food, blood samples were taken from all surviving animals and sent to the Sponsor for possible analysis for antibodies to T-AMG.

Blood samples (2.0 mL) were taken from the retro-orbital sinus, with the animals held under anaesthesia induced by isoflurane, into plain glass tubes and permitted to clot. The blood samples were separated at approximately 3000 rpm for 10 minutes and the serum stored at approximately -20°C, prior to despatch to the Sponsor.

The reporting of the results of any analysis is the responsibility of the Sponsor.

NECROPSY AND HISTOLOGY

Method of kill

The one animal that was killed during the study and those animals surviving until the end of the scheduled treatment period were killed by carbon dioxide asphyxiation. The sequence in which the animals were killed after completion of treatment was selected to allow satisfactory inter-group comparison.

Macroscopic pathology

All animals were subject to a detailed necropsy.

After a review of the history of each animal, a full macroscopic examination of the tissues was performed. All external features and orifices were examined visually. The cranial roof was removed to allow observation of the brain, pituitary gland and cranial nerves. After ventral mid-line incision, the neck and associated tissues and the thoracic, abdominal and pelvic cavities and their viscera were exposed and examined *in situ*. Any abnormal position, morphology or interaction was recorded.

The requisite organs were weighed and external and cut surfaces of the organs and tissues were examined as appropriate. Any abnormality in the appearance or size of any organ and tissue was recorded and the required tissue samples preserved in appropriate fixative.

The retained tissues were checked before disposal of the carcass.

Organ weights

The following organs, taken from each animal killed after 13 weeks of treatment, were dissected free of adjacent fat and other contiguous tissue and the weights recorded:

Adrenals
Brain
Epididymides
Heart
Kidneys
Liver
Ovaries
Spleen
Testes
Thymus
Uterus with cervix

Bilateral organs were weighed together.

Fixation

Testes and epididymides were fixed in Bouin's solution prior to transfer to 70% industrial methylated spirit and eyes were fixed in Davidson's fluid. Samples (or the whole) of the other tissues listed below from all animals were preserved in 10% neutral buffered formalin:

Adrenals	Ovaries
Aorta - thoracic	Pancreas
Brain	Pituitary
Caecum	Prostate
Colon	Rectum
Duodenum	Salivary glands - submandibular
Epididymides	- sublingual
Eyes#	Sciatic nerves+
Femurs+	Seminal vesicles
Head#	Skin
Heart	Spinal cord
Ileum	Spleen
Jejunum	Sternum
Kidneys	Stomach
Liver	Testes
Lungs	Thymus
Lymph nodes - mandibular	Thyroid with parathyroids
- mesenteric	Trachea
Mammary area	Urinary bladder
Oesophagus	Uterus and cervix

\+ Only one processed for examination
\# Not processed for examination

Samples of any abnormal tissues were also retained and processed for examination. In those cases where a lesion was not clearly delineated, contiguous tissue was fixed with the grossly affected region and sectioned as appropriate.

Samples of the head (including nasal cavity, paranasal sinuses and nasopharynx), eyes and the remaining femur and sciatic nerve were not examined histologically, but are retained against any future requirement for microscopic examination.

Histology

For those animals specified in the Pathology section, the relevant tissues were subject to histological processing.

Tissue samples were dehydrated, embedded in paraffin wax, sectioned at approximately four to five micron thickness and stained with haematoxylin and eosin.

Those tissues subject to histological processing included the following regions:

Adrenals	-	cortex and medulla
Brain	-	cerebellum, cerebrum and midbrain
Femur with joint	-	longitudinal section including articular surface, epiphysial plate and bone marrow
Heart	-	included auricular and ventricular regions.
Ileum	-	included peyers patches
Kidneys	-	included cortex, medulla and papilla regions
Liver	-	section from all main lobes
Lungs	-	section from two major lobes, to include bronchi
Salivary glands	-	submandibular and sublingual
Spinal cord	-	transverse and longitudinal section at the cervical, lumbar and thoracic levels
Sternum	-	included bone marrow
Stomach	-	included keratinised, glandular and antrum in sections
Thyroid	-	included parathyroids in section where possible
Uterus	-	uterus section separate from cervix section

For bilateral organs, sections of both organs were prepared. A single section was prepared from each of the remaining tissues required for microscopic pathology.

PATHOLOGY

Light microscopy

Microscopic examination was performed as follows:

All tissues preserved for examination (as specified above) were examined for all animals of Groups 1 (Control) and 4 (10.0 mL/kg/day) sacrificed on completion of the scheduled treatment period and for the animal killed during the study.

The adrenal glands of males in Group 2 (1.0 mL/kg/day) and 3 (3.3 mL/kg/day) were examined.

Tissues reported at macroscopic examination as being grossly abnormal were examined for all animals.

Findings were either reported as "present" or assigned a severity grade. In the latter case one of the following five grades was used - minimal, slight, moderate, marked or severe. A reviewing pathologist undertook a peer review of the microscopic findings.

DATA TREATMENT

This report contains serial observations pertaining to all weeks of treatment completed, together with signs data collected during the necropsy period. The only serial observations relating to the acclimatisation period included in this report relate to the ophthalmic examination.

Summary statistics (e.g. means and standard deviations) presented in this report were calculated from computer-stored individual raw data. The summary statistics and the individual data were stored in the computer to a certain number of decimal places, different for each parameter. For presentation purposes, however, they were usually rounded to fewer places. It is, therefore, not generally possible to reproduce the presented means and standard deviations exactly using the presented individual data.

The death codes in the appendices have the following meaning:

T	Terminal kill
W	Killed for welfare reasons

Throughout the report the following abbreviations are used:

N [or n]	Number of animals examined
SD [or sd]	Standard deviation

Definition of "Week"

The first week of treatment started at midnight prior to treatment commencing and ended at midnight on the seventh day following. Subsequent experimental weeks of treatment were of the same duration.

Signs

A detailed history of individual animals that showed signs is presented in Appendix 2, as the weeks in which the specified sign was observed. Given that none of these signs was related to treatment no tabular presentation of the data has been presented in this report. Similarly, there were no signs recorded associated with the dosing procedure and therefore no such signs are presented in this report.

Bodyweight

Analysis of variance was used on Week 13 absolute bodyweight values using Day 0 (Week 0) bodyweight as covariate. The adjusted Week 13 bodyweights are presented in Table 3.

Food consumption

Overall mean food consumption values were calculated from the weekly group mean values presented.

Values presented for the amount of food consumed in each cage in each experimental week allow for any animal that died or was killed during the week.

Weekly group mean food consumptions and standard deviations were derived from unrounded cage values, which were weighted to allow for any deaths.

Food conversion efficiency

Food conversion efficiency is the efficiency of conversion of food to new body tissue.

Group mean food conversion efficiencies were calculated for each week of treatment, and were derived from unrounded cage values.

The weekly group mean values presented were calculated from the individual food consumption values and the individual bodyweight gain of animals alive at the end of the week.

$$\text{Food conversion efficiency} \quad \frac{\text{Bodyweight gain (g)}}{\text{Total food consumed (g/animal)}} \times 100$$

Overall group mean values were calculated as the overall group mean bodyweight gain, divided by the total food consumption, expressed as a percentage.

Water consumption

Water consumption was calculated from measurements of initial and final weights (g) of the water bottle and contents for each cage (it was assumed that 1 mL of water weighed 1 g).

Ophthalmic examination

All observations made at ophthalmoscopic examination are recorded in the raw data. This report only contains those that were considered to be unusual or abnormal and those which may correlate with findings recorded elsewhere in this report, for example those observed at micropathology. Observations were bilateral unless otherwise indicated.

Haematology

The abbreviation used in Appendix 9 has the following meaning:

CTD Clotted sample

Blood chemistry

Albumin to globulin (A/G) ratios were calculated as:

$$\text{A/G} \quad \frac{\text{Albumin concentration}}{\text{Total protein albumin concentration}}$$

Organ weights

The adjusted organ weights presented in Table 9 and Appendix 11 were those from the analysis of covariance. This analysis took the transformed absolute organ weights as the response variable and the original, untransformed, terminal bodyweights as the covariate. Linear regression lines were fitted to each group, assuming the lines for all groups were parallel to one another. The individual organ weights (and therefore the group mean organ weights) were adjusted towards the overall mean terminal bodyweight, thus removing the effect of bodyweight. The standard deviations were based on the variability after allowing the covariate, rather than the usual unadjusted standard deviations.

Pathology

Tissues which could not be examined are specified in Appendix 12. The absence of a comment for a tissue scheduled for examination therefore indicates that the tissue was examined and found to be normal. Tissues recorded as abnormal macroscopically but found to be normal microscopically are described as 'No significant lesion' in the microscopic pathology appendix. In all tabular presentations of data the tissues specified in the protocol for histopathological examination precede other tissues.

Statistical analysis

Standard deviations were calculated as considered appropriate.

Statistical evaluation of grip strength, motor activity, bodyweight, haematology, blood chemistry, organ weights and any data derived from these was performed according to the following schedule:

1.	Outliers to be excluded (reason indicated)		
2.	Test homogeneity of variance (Bartlett's test)	Significant : Not significant:	Go to 3. Go to 6 or 7.
3.	Transformation log→2→	Significant : Not significant:	Go to 4. Go to 6 or 7.
4.	Transformation 1/ x→2→	Significant : Not significant:	Go to 5. Go to 6 or 7.
5.	Transformation√ x→2→	Significant : Not significant:	Go to 8. Go to 6 or 7.

6. Analysis of covariance (ANOCOVA). Test slope of linear regression: if significant ANOCOVA was completed and if significant a Student's t-test was performed; then the test was terminated. If not significant go to 7. This test was used for the Week 13 bodyweight, with the bodyweight on Day 0 (Week 0) as covariate, and organ weights with the autopsy bodyweight as covariate.

7. Analysis of variance (ANOVA); if significant a Student's t-test was performed; test terminated. This test was used for all other individual numerical data.

8. Kruskal-Wallis; if significant a Steel's test was performed; test terminated. For convenience, the result of this test for organ weight is presented in the appendix of adjusted organ weights.

All statistical analyses, were performed on unrounded values. The method of each analysis is indicated on the individual appendix for Controls. The method is indicated as a one or two figure code, e.g. 2.7 means a non-significant results from the homogeneity of variance (2) followed by ANOVA and Student's t-test (7).

Categorical data were analysed, where considered appropriate, using the Fisher's Exact Probability test.

Unless stated, group mean values or incidences for the treated groups were not significantly different from those of the Controls ($p>0.05$).

Significant differences between Control and treated groups were expressed at the 5% ($p<0.05$) or 1% ($p<0.01$) level. The following statistical cyphers were used throughout the report:

a - $p<0.05$; b - $p<0.01$	- using categorical or parametric tests
A - $p<0.05$; B - $p<0.01$	- using non-parametric tests
n.s	- not significant

QUALITY ASSURANCE AND ARCHIVING PROCEDURES

Quality Assurance

Details of the Quality Assurance inspections and audits (undertaken at Huntingdon Life Sciences) are presented on the Quality Assurance Statement.

The formulation analyses data and reporting were subject to Quality Assurance inspections and audits at the Principal Investigator's facility. A separate Quality Assurance Statement is presented in this report (Annex 1).

Archives

Following completion of this study all raw data, specimens and samples, except those generated or used during any Sponsor's or supplier's analysis, were stored in the archives of Huntingdon Life Sciences. Types of sample and specimen which are unsuitable, by reason of instability, for long term retention and archiving may be disposed of after the periods stated in Huntingdon Life Sciences Standard Operating Procedures.

A copy of the final report and all Quality Assurance inspection records will be retained indefinitely. All other appropriate specimens and records will be retained for a minimum period of five years from the date of issue of the final report. At the end of the five year retention period the Sponsor will be contacted and advice sought on the above requirements. Under no circumstances will any item be discarded without the Sponsor's knowledge.

All documentation, samples and data pertaining to the analytical aspect of the study undertaken by the Principal Investigator will be transferred to the Sponsor's archive for storage.

DEVIATIONS FROM PROTOCOL

The following deviations from protocol occurred:

> On one occasion during the treatment period (20 January 2006) the relative humidity in the animal room was low (35%), but this did not result in any overt changes in the animals.

This deviation was considered to have not affected the integrity or validity of the study.

UNEXPECTED EVENTS

Upon receipt at the Safety and Toxicology Department of Novozymes A/S, the formulation samples for Week 1, 6 and 13 were stored, as scheduled, in a freezer. They were then stored in a freezer within the Enzyme Analytical Laboratory, but this freezer subsequently malfunctioned and the samples thawed. The samples were analyzed as planned and the results were as expected, in terms of enzyme activity. Consequently, this unexpected event was not considered to have affected the validity of the study.

RESULTS

Formulation chemistry (Annex 1)

The analysis of test formulations prepared in Week 1, 6 and 13 indicated that the mean achieved enzyme activities were between 98.7 to 102% of intended, demonstrating satisfactory formulation. The test enzyme was not detected in the Control formulations (<0.21 AGU/g). In addition, there was no significant difference between the Tox-batch of the enzyme (Batch PPY24900) and the highest dosage formulation.

Mortality (Appendix 1)

There were no deaths during the treatment period that were considered related to the administration of the test material.

One male receiving 3.3 mL/kg/day (Animal No. 10) was killed for welfare reasons during Week 3. In Week 3 this animal was reported as being underactive with partially closed eyelids and on the day after these signs were first apparent the condition of the animal deteriorated. Signs included underactive behaviour, body tremor, abnormal gait and brown staining around the eyes/nose and, consequently, the animal was killed. The cause of death was identified as meningitis in the brain and spinal cord and, consequently, this death was considered incidental and not related to treatment.

Detailed physical examination and arena observations (Appendix 2)

Routine weekly clinical observations did not indicate any changes in the appearance or general behaviour of the animals that were considered to be related to treatment.

Sensory reactivity observations and grip strength (Table 1; Appendix 3)

Sensory reactivity was unaffected by treatment.

Forelimb grip strength values at all doses in males were lower than those for Controls ($p<0.01$) but there was no dose-relationship. The mean forelimb value for concurrent Control males (1.43 kg) exceeded the historical control range (0.62-1.33 kg; mean 0.96 kg; 149 animals) whilst that for the treated groups (1.18, 1.26 or 1.21 kg for males receiving 1.0, 3.3 or 10.0 mL/kg/day, respectively) were within this range. Consequently, these differences from controls were not attributed to treatment.

Historical control data for grip strength in male CD rats

Study code	Date	No. of animals	Grip strength (kg)	Study code	Date	No. of animals	Grip strength (kg)
A	Jul-03	10	0.63	I	Nov-04	9	0.84
B	Aug-03	10	0.62	J	Jan-05	10	1.11
C	Nov-03	10	1.18	K	Mar-05	10	0.96
D	Dec-03	10	1.21	L	Mar-05	10	1.33
E	Apr-04	10	1.21	M	May-05	10	0.66
F	Jun-04	10	0.83	N	Jun-05	10	1.04
G	Jul-04	10	0.87	O	Nov-05	10	1.24
H	Sep-04	10	0.71				

Motor activity (Figure 1; Table 2; Appendix 4)

There were no toxicologically significant findings at the motor activity investigation in Week 12.

Low beam scores (cage floor activity) and, to a lesser extent, high beam scores (rearing activity) for males receiving 10 mL/kg/day were higher than those of Controls from 18 minutes to the end of the one-hour recording period. Only one time interval value (low beam score at 48 minutes) achieved statistical significance, but five of the ten time interval low beam values exceeded the maximum in the historical control range, although the total low beam score was just within this range (see Table 2 for historical control range). In addition, an examination of individual data revealed that the Controls generally had a higher incidence of low (<10) or zero values after the first 18 minutes of the recording period. There was no similar finding in the females and these animals, conversely, showed a small reduction of low and high beam breaks. Consequently, the variations seen in males are considered unlikely to be of any toxicological significance.

Bodyweight (Figure 2; Table 3; Appendix 5)

Overall bodyweight gain was considered to be unaffected by treatment; the Week 13 adjusted bodyweight values were not statistically significantly different from the Controls.

Bodyweight gains, when compared with the Controls, were slightly low during the first week of treatment for females receiving 1.0 mL/kg/day (0.88xControl), males receiving 3.3 mL/kg/day (0.90xControl) and animals receiving 10 mL/kg/day (up to 0.85xControl). Bodyweight gains were also slightly low (0.87xControl) during the second week of treatment for females receiving 1.0 or 3.3 mL/kg/day. In general, these differences did not show any clear dose-relationship and as overall weight gain was not affected, these differences during the early part of the treatment period were considered unlikely to be of any toxicological significance.

Food consumption (Table 4; Appendix 6)

There was no toxicologically significant effect upon food consumption.

Food consumption was consistently slightly low for males receiving 10.0 mL/kg/day, resulting in an overall food consumption that was approximately 94xControl. The difference from Controls was, however, small and was confined to one cage (Cage 5) and, consequently, was considered unlikely to be of any toxicological significance. Females at this dosage were unaffected, but those receiving 3.3 mL/kg/day also showed a slight overall reduction of food intake (0.93xControl). Females receiving 10.0 mL/kg/day were clearly not affected and the difference from Controls in the intermediate dose females was not, therefore, attributed to treatment.

Food conversion efficiency (Table 5)

Food conversion efficiency was not affected by treatment.

Water consumption (Table 6; Appendix 7)

There was no effect of treatment upon water consumption.

A number of small differences from controls occurred during the treatment period but these were inconsistent or lacked dosage-relationship and were attributed to normal biological variation.

Ophthalmic examination (Appendix 8)

Ophthalmic investigations performed in Week 13 did not reveal any changes that were considered treatment-related.

Haematology (Table 7; Appendix 9)

Haematology investigations performed during Week 13 did not reveal any findings that were considered to be toxicologically significant.

Haematocrit for males receiving 10.0 mL/kg/day was statistically significantly low when compared with the Controls but the difference from Control (0.95x) was minor, all individual values were within the normal range (Range 0.416-0.492 L/L; mean±sd 0.453±0.023; number of values 305) and no similar effect was seen in any other erythrocytic parameter, or in females. Consequently, this was attributed to normal biological variation.

Blood chemistry (Table 8; Appendix 10)

There were no toxicologically significant changes in the blood plasma in Week 13, with all inter-group differences being small or lacking dose-relationship.

High plasma glucose concentrations (up to 1.16xControl) were reported in all treated groups of males. The differences from Control were not dose-related and as all individual values were within the background range (Range 5.29-9.89 mmol/L; mean±sd 7.42±1.45; number of values 304) and there was no similar change in females, this was considered not to be of toxicological significance.

Creatinine concentrations were statistically significantly high ($p<0.01$; up to 1.23xControl) for all treated groups of females; males were unaffected. An examination of the individual values indicated that the majority of values, including those of the Controls, were below the lower limit of the background range (Range 47-62 mmol/L; mean±sd 54±5; number of values 291) and in view of this, the variations of creatinine were not considered to be related to treatment.

Plasma calcium concentrations were slightly low (0.97xControl) for females receiving 3.3 or 10.0 mL/kg/day. There was, however, no difference in the magnitude of change for females receiving these doses and no similar trend was seen in males. In view of the lack of dose-relationship and the small magnitude of change, this was considered unlikely to be related to treatment.

Plasma albumin concentrations were slightly low ($p<0.05$; 0.92xControl) in males receiving 10.0 mL/kg/day but 9/10 individual values were normal (Range 30-38 g/L; mean±sd 35±2; number of values 192) and this slight change was considered to be of no toxicological significance.

Organ weights (Table 9; Appendix 11)

Organ weights were unaffected by treatment.

Macropathology (Table 10; Appendix 12)

There were no macroscopic findings that were considered related to treatment. All changes were of the type encountered normally in young CD rats at these laboratories.

Histopathology (Table 11; Appendix 12)

Adrenals

When compared with the Controls, there was an increased incidence or increased severity of cortical vacuolation in males receiving 3.3 or 10.0 mL/kg/day.

Treatment-related microscopic findings in the adrenals

Group/sex		1M	2M	3M	4M
Dose (mL/kg/day)		0	1.0	3.3	10.0
Cortical vacuolation	Total	2	2	4	5
	Minimal	2	2	3	4
	Slight	0	0	1	1
Number of animals examined		10	10	10	10

All other microscopic findings were considered to be incidental and of no toxicological importance.

DISCUSSION

The oral administration of T-AMG, PPY 24900 (a food enzyme) to CD rats at doses up to 10 mL/kg/day (equivalent to 5528 AGU/kg/day or 1.47 gTOS/kg/day) for 13-weeks was well tolerated, with no treatment-related deaths of signs of toxicity. There was no effect upon weight gain and food and water intake, nor was there any change in the blood that could, with any confidence, be attributed to treatment.

The only finding that was possibly due to treatment was a small increase in the incidence and severity of cortical vacuolation in the adrenal glands of males. Since glucocorticoids secreted by the adrenal cortex are involved in, amongst other functions, protein catabolism it is possible that this could be an adaptive response to the administration of the T-AMG enzyme. Some vacuolation of the adrenal cortex at a minimal level was reported in two Control males. This suggested that the slightly increased incidence (4/10 males at 3.3 mL/kg/day and 5/10 males at 10 mL/kg/day) and severity (one male at each of these two doses were reported as having slight vacuolation with all other affected animals showing only minimal change) may represent a mild exacerbation of a normal finding. The incidence of this change (20% in Controls, 20% in low dose animals, 40% in intermediate dose animals and 50% in high dose animals) was generally within the ranges observed in other 13-week toxicity studies in Control CD rat (Study range 0 to 42% from seven recent studies; overall incidence (16/71 (22.5%)) and the finding at slight severity was present in 6/71 animals (8.5%)). There was no evidence for any other degenerative change, such as inflammation or necrosis, and females were not affected. In view of this, the findings in the adrenal of males were considered unlikely to be of any toxicological significance.

In view of the absence of any toxicologically significant change in this study, the no-observed-adverse-effect level (NOAEL) was considered to be 10 mL/kg/day, which is equivalent to 5528 AGU/kg/day or 1.47 gTOS/kg/day.

CONCLUSION

It is concluded that oral administration of T-AMG, PPY 24900 (a food enzyme) to CD rats at doses up to 10 mL/kg/day for 13 weeks was well tolerated and did not cause any toxicologically significant change. In males receiving 3.3 or 10 mL/kg/day there was a small increased in incidence and severity of adrenal cortical vacuolation but this was not considered toxicologically significance. Consequently, the no-observed-adverse-effect level (NOAEL) in this study was considered to be 10 mL/kg/day, which is equivalent to 5528 AGU/kg/day or 1.47 gTOS/kg/day.

REFERENCES

ANGERVALL, L. and CARLSTROM, E. (1963) Theoretical criteria for the use of relative organ weights and similar ratios in biology. *J Theoret. Biol*., **4**, 254-259.

ARMITAGE, P. (1971) *Statistical Methods in Medical Research,* p.189. Blackwell Scientific Publications, Oxford, UK.

BARTLETT, M.S. (1937) Properties of sufficiency and statistical tests. *Proceedings of the Royal Society. Series A*, **160**, 268-282. (Also see J.R.Stat.Soc.Suppl 4,1937, published later)

FISHER, R.A. (1973) *Statistical Methods for Research Workers,* 14th edn., p.96. Hafner Publishing Company, New York, USA.

KRUSKAL, W.H. and WALLIS, W.A. (1952) Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, **47**, 583-621.

KRUSKAL, W.H. and WALLIS, W.A. (1953) Errata for Kruskal-Wallis(1952). *Journal of the American Statistical Association*, **48**, 907-911

STEEL, R.G.D. (1959) A multiple comparison rank sum test: treatments versus control, *Biometrics*, **15**, 560-572.

FIGURE 1

Motor activity - group mean scores (beam breaks) for males during Week 12 of treatment

Group 1: Control
Group 2: T-AMG, PPY 24900: 1.0 mL/kg/day
Group 3: T-AMG, PPY 24900: 3.3 mL/kg/day
Group 4: T-AMG, PPY 24900: 10.0 mL/kg/day

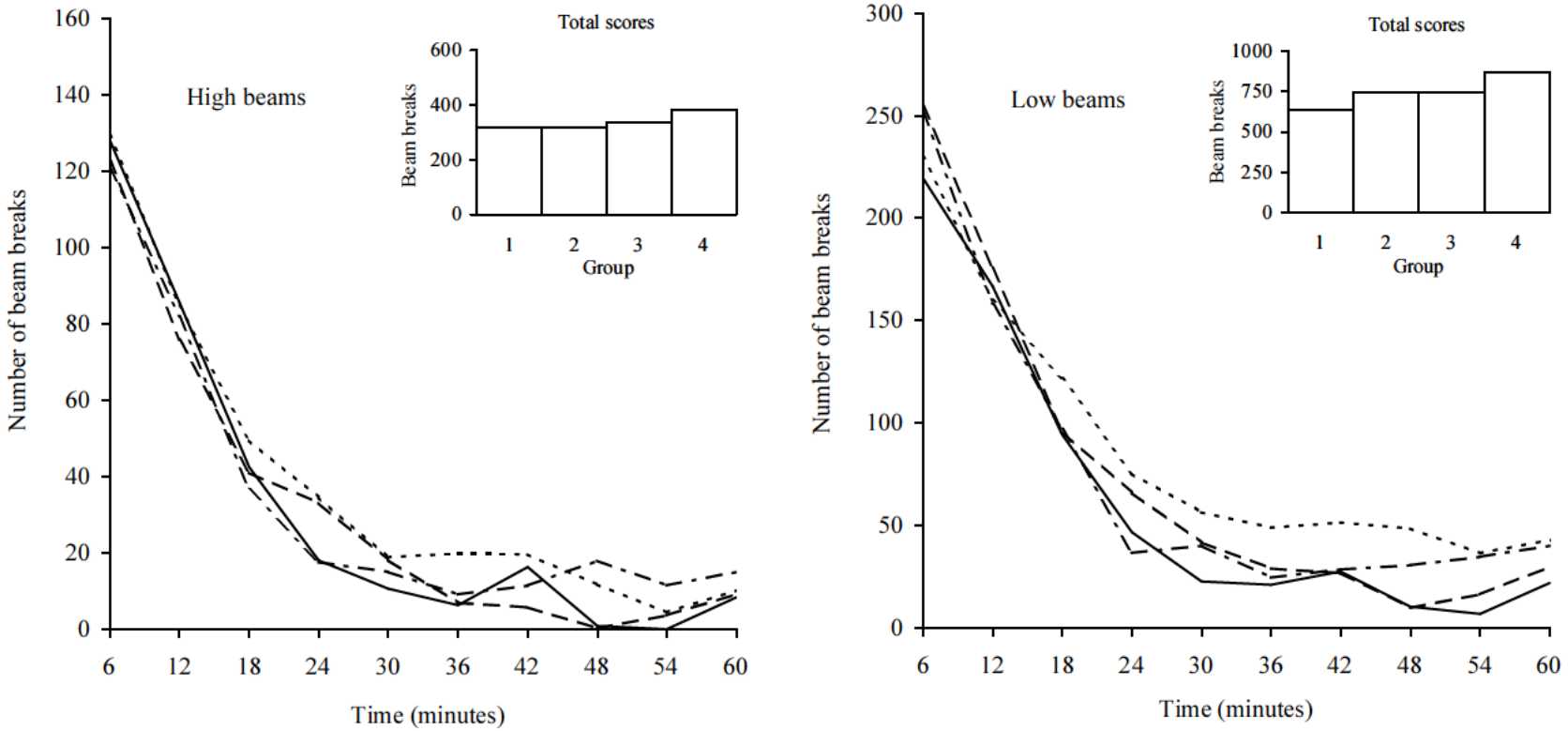
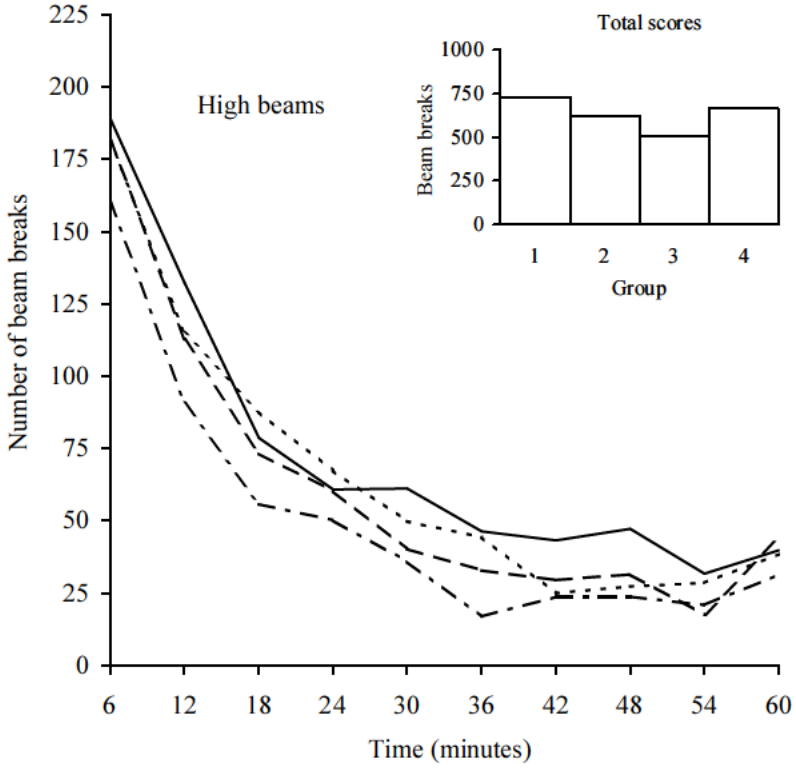


FIGURE 1 - continued

Motor activity - group mean scores (beam breaks) for females during Week 12 of treatment

Group 1: Control
Group 2: T-AMG, PPY 24900: 1.0 mL/kg/day
Group 3: T-AMG, PPY 24900: 3.3 mL/kg/day
Group 4: T-AMG, PPY 24900: 10.0 mL/kg/day



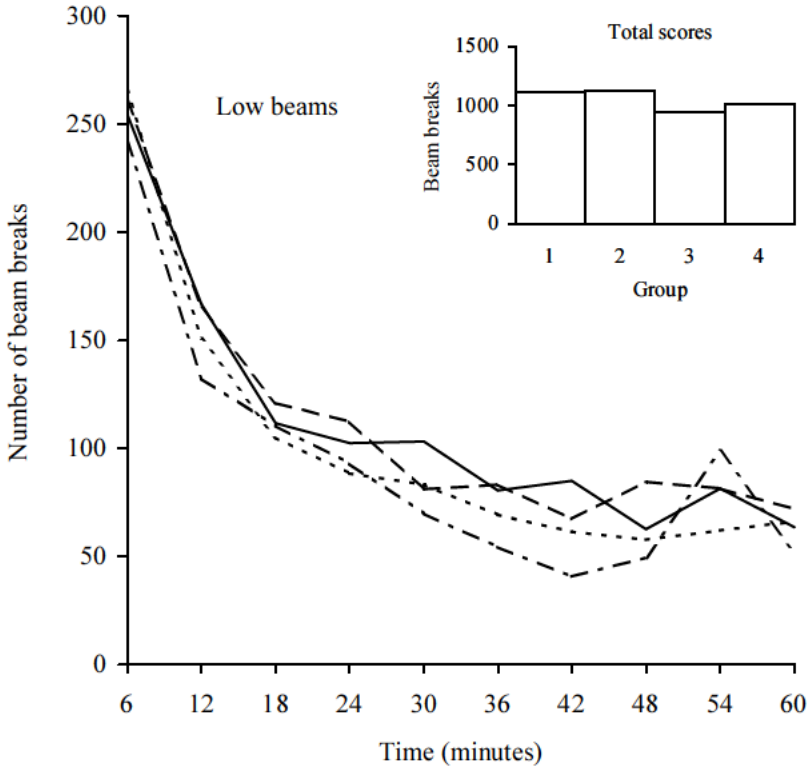


FIGURE 2

Bodyweight - group mean bodyweight versus period of treatment - males

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

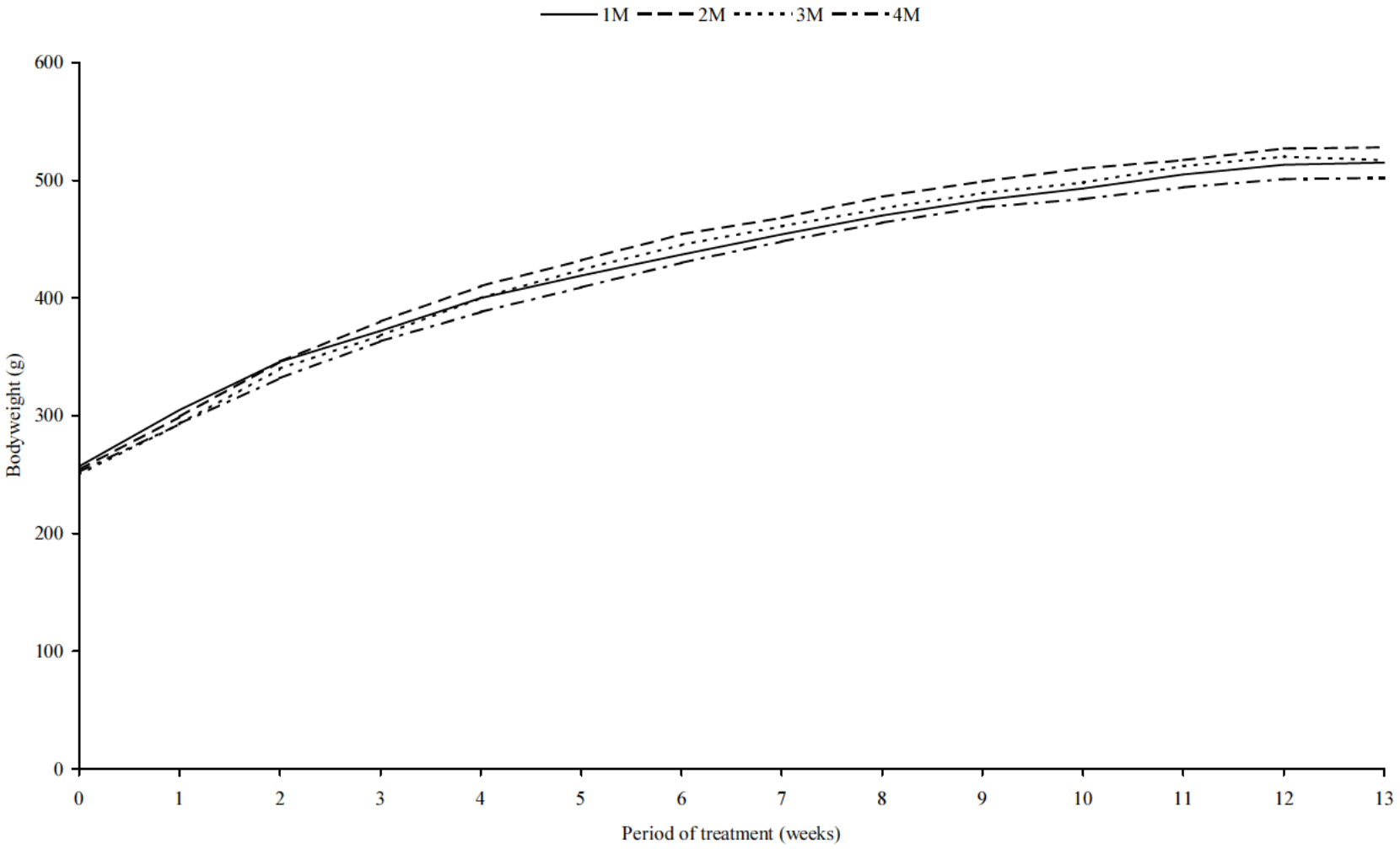


FIGURE 2 - continued

Group mean bodyweight versus period of treatment - females

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

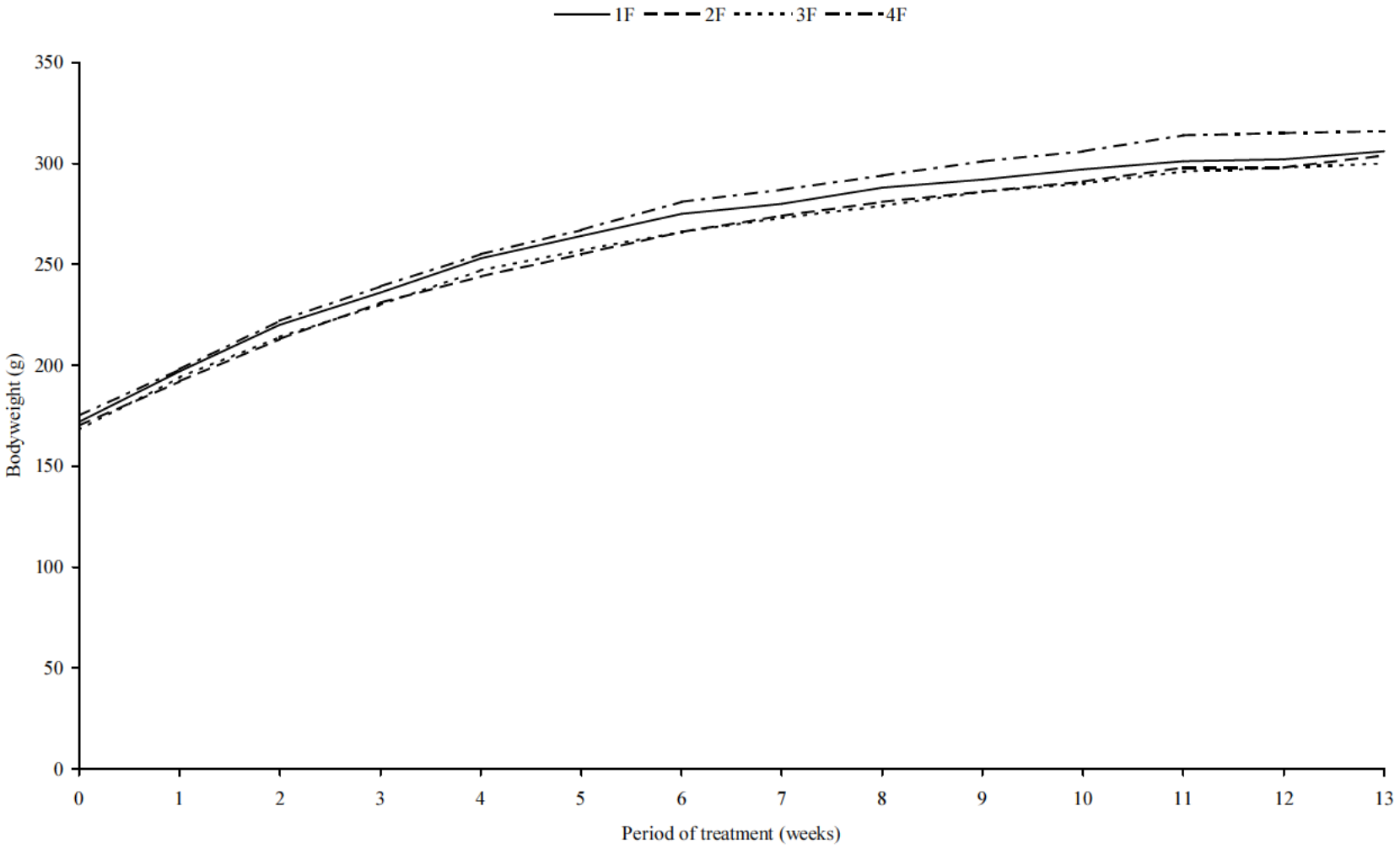


TABLE 1

Sensory reactivity and grip strength - summary of findings during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group/sex:		1M	2M	3M	4M	1F	2F	3F	4F
Number of animals:		10	10	9	10	10	10	10	10
Parameter	Grade								
Approach response (1-3)	2	10	10	9	10	10	10	10	10
Touch response (1-3)	2	10	10	9	10	10	10	10	10
Auditory startle reflex	2	3	0	0	0	1	0	2	1
(1-4)	3	7	10	9	10	9	10	8	9
Tail pinch response	2	8	1	2	1	6	3	2	2
(1-4)	3	2	9	6	7	4	7	8	8
	4	0	0	1	2	0	0	0	0
Forelimb grip strength (kg)	Mean	1.43	1.18 b	1.26 b	1.21 b	1.11	1.10	1.07	1.04
	SD	0.10	0.15	0.08	0.12	0.12	0.16	0.12	0.13
Hindlimb grip strength (kg)	Mean	0.60	0.61	0.60	0.60	0.48	0.45	0.49	0.47
	SD	0.10	0.10	0.07	0.07	0.08	0.07	0.05	0.04

Significant when compared with Group 1: a - $p<0.05$; b - $p<0.01$

TABLE 2

Motor activity - group mean scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Number of animals	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
1M	10	High	128.2	85.6	42.5	18.0	10.6	6.3	16.3	0.9	0.0	8.3	316.7
		SD	25.3	34.3	29.0	28.0	18.8	9.7	29.4	1.5	0.0	13.4	127.8
2M	10	High	123.1	75.8	41.0	33.1	18.2	6.9	5.8	0.2	3.6	9.2	316.9
		SD	14.9	20.0	21.3	23.6	24.6	12.0	12.2	0.6	9.2	19.9	97.4
3M	9	High	121.0	82.0	36.6	17.6	15.0	9.1	11.3	17.9	11.4	15.0	336.9
		SD	28.5	37.3	23.1	24.8	20.4	14.5	22.1	26.5	19.4	16.6	172.8
4M	10	High	129.7	84.6	49.5	34.5	18.9	19.8	19.6	11.7	4.4	10.1	382.8
		SD	28.5	28.9	26.2	32.2	20.4	22.3	24.5	19.4	9.4	20.3	150.4
1M	10	Low	219.6	166.7	94.6	46.7	22.7	21.1	27.5	10.5	6.9	21.9	638.2
		SD	31.4	51.3	34.9	39.4	37.0	31.7	38.5	14.0	13.2	28.9	188.4
2M	10	Low	255.2	175.8	95.9	66.1	41.7	28.9	26.8	9.8	16.3	29.8	746.3
		SD	42.6	41.7	49.4	42.6	38.3	34.1	40.3	12.5	31.1	59.1	211.9
3M	9	Low	252.2	158.2	97.0	36.6	40.1	24.6	28.6	30.6	34.7	40.1	742.6
		SD	43.9	47.0	46.4	32.9	44.1	32.1	31.9	38.5	41.8	44.4	241.1
4M	10	Low	230.3	159.9	121.8	75.3	56.4	48.9	51.4	48.5 A	36.5	43.0	872.0
		SD	45.0	44.5	33.3	29.8	36.6	45.6	34.0	42.0	36.4	52.0	150.9

Significant when compared with Group 1: A - $p<0.05$; B - $p<0.01$

TABLE 2 - continued

Motor activity - group mean scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Number of animals	Beam level	Time (minutes)										
			6	12	18	24	30	36	42	48	54	60	Total
1F	10	High	189.2	132.4	78.7	60.8	61.1	46.4	43.3	47.3	31.7	39.8	730.7
		SD	40.1	44.4	42.9	32.4	30.2	31.2	41.2	36.2	24.5	33.6	243.8
2F	10	High	182.0	112.7	73.2	60.2	40.3	32.8	29.5	31.5	17.5	44.0	623.7
		SD	44.1	54.8	33.7	43.7	25.8	37.9	23.0	29.9	10.4	32.1	267.4
3F	10	High	160.4	90.6	55.7	50.1	36.0	16.9	23.7	23.7	20.9	31.4	509.4
		SD	35.6	44.4	32.4	29.0	29.6	24.5	32.9	34.7	13.6	36.7	255.3
4F	10	High	181.4	114.9	87.7	67.3	49.8	44.4	25.0	27.3	28.5	38.5	664.8
		SD	33.8	55.4	50.6	40.6	38.9	36.8	20.7	27.4	17.6	25.2	255.9
1F	10	Low	254.6	166.6	111.7	102.4	103.2	80.4	84.9	62.7	81.4	63.6	1111.5
		SD	54.5	42.9	29.8	51.5	43.2	42.6	65.6	40.5	65.2	41.4	204.0
2F	10	Low	261.2	164.7	120.8	112.3	81.1	83.2	67.3	84.4	81.4	71.9	1128.3
		SD	67.1	72.8	47.7	58.3	34.3	41.1	44.1	35.7	67.0	38.1	251.3
3F	10	Low	242.2	132.3	110.5	92.9	69.7	54.3	40.7	49.5	98.7	51.6	942.4
		SD	24.1	30.5	35.8	42.2	45.1	56.5	37.4	40.4	119.4	29.7	306.9
4F	10	Low	266.0	150.7	104.9	88.3	83.2	69.2	61.5	57.8	62.0	66.2	1009.8
		SD	59.6	56.1	61.7	63.2	42.9	28.5	40.2	51.6	42.8	41.1	319.5

$p \geq 0.05$, no statistical significance

TABLE 2 - continued

Motor Activity - Historical control data for Crl: CD(SD) male rats for six studies performed at Huntingdon Life Sciences; Week 12 data

Study code	Date	Number of animals	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
A	Jul-03	10	High	121.2	73.8	43.8	45.2	27.4	24.9	20.2	12.8	9.3	13.4	392.0
B	Aug-03	10	High	95.6	72.3	40.8	21.9	18.9	16.3	11.5	10.8	13.7	21.2	323.0
C	Jun-04	10	High	125.6	73.3	43.7	22.7	11.5	12.0	10.9	3.1	0.3	5.5	308.6
D	Sep-04	10	High	103.3	45.8	23.8	15.9	4.9	2.7	8.2	4.1	5.2	2.4	216.3
E	Mar-05	10	High	153.1	99.2	59.4	33.8	39.5	29.2	36.3	27.8	35.9	20.5	534.7
F	Nov-05	10	High	108.3	58.8	25.1	24.5	15.9	9.4	8.9	15.4	8.9	4.7	279.9
	Total of 6 studies													
	Mean			117.9	70.5	39.4	27.3	19.7	15.8	16.0	12.3	12.2	11.3	342.4
	Minimum			95.6	45.8	23.8	15.9	4.9	2.7	8.2	3.1	0.3	2.4	216.3
	Maximum			153.1	99.2	59.4	45.2	39.5	29.2	36.3	27.8	35.9	21.2	534.7

Study code	Date	Number of animals	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
A	Jul-03	10	Low	157.9	125.6	94.2	62.0	53.6	53.1	35.7	37.8	43.2	37.9	701.0
B	Aug-03	10	Low	210.7	155.0	96.8	51.2	50.5	30.1	40.0	27.6	25.4	55.4	742.7
C	Jun-04	10	Low	218.7	144.4	72.4	52.6	38.1	26.3	23.7	19.7	4.6	19.5	620.0
D	Sep-04	10	Low	204.0	100.8	76.7	48.0	25.7	21.0	31.7	24.4	32.3	21.6	586.2
E	Mar-05	10	Low	241.6	153.9	114.8	60.3	64.9	46.4	43.9	48.3	52.4	47.9	874.4
F	Nov-05	10	Low	188.9	123.0	72.6	72.6	53.2	31.7	43.6	27.9	27.3	29.3	670.1
	Total of 6 studies													
	Mean			203.6	133.8	87.9	57.8	47.7	34.8	36.4	31.0	30.9	35.3	699.1
	Minimum			157.9	100.8	72.4	48.0	25.7	21.0	23.7	19.7	4.6	19.5	586.2
	Maximum			241.6	155.0	114.8	72.6	64.9	53.1	43.9	48.3	52.4	55.4	874.4

TABLE 3

Bodyweight - group mean values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Statistic	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10
1M	Mean	257	305	346	372	400	419	437	454	470	483	493
	SD	14.8	22.7	27.3	33.9	39.0	40.2	44.2	51.7	58.1	62.3	65.5
	N	10	10	10	10	10	10	10	10	10	10	10
2M	Mean	254	299	346	380	410	432	454	468	486	499	510
	SD	13.7	18.8	25.4	32.6	39.1	42.8	47.0	46.1	47.2	48.0	53.6
	N	10	10	10	10	10	10	10	10	10	10	10
3M	Mean	250	293	340	368	400	424	445	461	476	489	498
	SD	12.4	17.2	26.6	30.4	39.3	40.9	44.4	48.0	48.1	51.5	56.2
	N	10	10	10	9	9	9	9	9	9	9	9
4M	Mean	252	293	332	363	388	409	430	448	464	477	484
	SD	10.1	14.0	22.6	28.7	33.1	38.6	43.5	47.8	50.0	52.5	56.7
	N	10	10	10	10	10	10	10	10	10	10	10

$p \geq 0.05$, no statistical significance

TABLE 3 - continued

Bodyweight - group mean values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Statistic	Week 11	Week 12	Week 13	Week 13 adjusted values
1M	Mean	505	513	515	502
	SD	67.2	68.1	66.1	66.1
	N	10	10	10	10
	Method †				2.6
	P				0.496
2M	Mean	517	527	528	525
	SD	52.7	58.3	54.8	54.8
	N	10	10	10	10
3M	Mean	512	520	517	531
	SD	58.0	59.8	58.9	58.9
	N	9	9	9	9
4M	Mean	494	501	502	506
	SD	60.0	60.1	58.4	58.4
	N	10	10	10	10

† See Statistical Analysis Section in Experimental Procedure for explanation
p≥0.05, no statistical significance

TABLE 3 - continued

Bodyweight - group mean values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Statistic	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10
1F	Mean	172	197	220	236	253	264	275	280	288	292	297
	SD	10.2	10.2	12.2	15.1	17.2	15.3	16.0	20.6	22.1	19.3	19.6
	N	10	10	10	10	10	10	10	10	10	10	10
2F	Mean	170	192	213	231	244	255	266	274	281	286	291
	SD	11.7	11.9	13.4	14.4	17.8	15.7	18.3	17.5	19.7	16.7	18.7
	N	10	10	10	10	10	10	10	10	10	10	10
3F	Mean	168	194	214	230	247	257	266	273	279	286	290
	SD	9.5	11.6	14.7	16.1	20.9	20.8	20.1	19.8	22.3	23.4	25.0
	N	10	10	10	10	10	10	10	10	10	10	10
4F	Mean	175	198	222	239	255	267	281	287	294	301	306
	SD	11.9	15.0	18.9	21.6	21.9	22.3	27.0	31.7	28.6	31.7	31.9
	N	10	10	10	10	10	10	10	10	10	10	10

$p \geq 0.05$, no statistical significance

TABLE 3 - continued

Bodyweight - group mean values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Statistic	Week 11	Week 12	Week 13	Week 13 adjusted values
1F	Mean	301	302	306	305
	SD	22.9	23.2	19.8	19.8
	N	10	10	10	10
	Method †				2.6
	P				0.919
2F	Mean	298	298	304	307
	SD	19.2	21.3	23.5	23.5
	N	10	10	10	10
3F	Mean	296	298	300	305
	SD	23.0	27.1	27.3	27.3
	N	10	10	10	10
4F	Mean	314	315	316	309
	SD	36.0	36.4	33.9	33.9
	N	10	10	10	10

† See Statistical Analysis Section in Experimental Procedure for explanation
p≥0.05, no statistical significance

TABLE 4

Food consumption - group mean values (g/animal/week)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12
1M	Mean	194	189	189	186	190	220	193	195	193	191	193	176
	SD	13.1	10.6	13.3	8.3	9.1	51.4	12.1	13.5	12.2	13.9	13.2	0.5
	N	2	2	2	2	2	2	2	2	2	2	2	2
2M	Mean	189	197	198	192	197	197	197	196	194	193	192	190
	SD	3.5	1.5	1.9	5.2	4.9	7.3	9.8	6.6	5.5	6.6	4.2	8.7
	N	2	2	2	2	2	2	2	2	2	2	2	2
3M	Mean	190	189	183	191	197	197	196	194	193	187	187	182
	SD	7.2	8.2	2.0	12.7	4.4	4.6	10.1	11.2	3.0	3.9	3.4	5.7
	N	2	2	2	2	2	2	2	2	2	2	2	2
4M	Mean	180	178	179	174	185	186	183	187	186	181	176	172
	SD	4.8	11.7	10.4	6.4	8.3	8.7	8.0	6.9	7.9	9.5	7.8	7.0
	N	2	2	2	2	2	2	2	2	2	2	2	2

TABLE 4 - continued

Food consumption - group mean values (g/animal/week)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Week 13	Mean 1-13	As % of Control
1M	Mean	182	192	-
	SD	11.7		
	N	2		
2M	Mean	178	193	101
	SD	0.1		
	N	2		
3M	Mean	177	189	98
	SD	4.9		
	N	2		
4M	Mean	166	180	94
	SD	5.1		
	N	2		

TABLE 4 - continued

Food consumption - group mean values (g/animal/week)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12
1F	Mean	126	133	136	139	142	146	142	135	135	132	133	129
	SD	0.6	1.6	1.2	4.8	6.8	5.7	3.9	4.7	5.5	3.8	4.8	4.9
	N	2	2	2	2	2	2	2	2	2	2	2	2
2F	Mean	123	128	138	137	140	145	142	137	136	135	139	133
	SD	5.4	5.1	7.2	7.4	3.0	8.3	2.2	5.3	2.4	10.6	3.7	8.1
	N	2	2	2	2	2	2	2	2	2	2	2	2
3F	Mean	120	123	129	131	130	130	131	126	128	124	127	121
	SD	2.9	1.5	2.3	0.6	4.6	2.7	2.3	1.0	1.9	3.2	0.1	3.5
	N	2	2	2	2	2	2	2	2	2	2	2	2
4F	Mean	127	131	139	140	147	147	147	141	138	139	143	132
	SD	1.5	3.8	3.6	1.0	0.8	3.5	3.6	5.0	3.3	7.3	6.1	2.7
	N	2	2	2	2	2	2	2	2	2	2	2	2

TABLE 4 - continued

Food consumption - group mean values (g/animal/week)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Week 13	Mean 1-13	As % of Control
1F	Mean	125	135	-
	SD	4.4		
	N	2		
2F	Mean	125	135	100
	SD	5.4		
	N	2		
3F	Mean	119	126	93
	SD	3.3		
	N	2		
4F	Mean	125	138	102
	SD	7.8		
	N	2		

TABLE 5

Food conversion efficiency - group mean values (%)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group/ Sex	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11
1M	24.5	21.4	13.8	15.4	9.7	8.4	8.6	8.3	6.5	5.3	6.5
2M	23.6	24.1	17.0	15.7	10.9	11.5	7.2	8.8	7.0	5.4	3.9
3M	22.4	24.9	17.7	16.8	12.2	10.7	8.1	7.8	6.6	4.9	7.5
4M	22.9	22.2	17.2	14.4	11.2	11.0	10.3	8.4	6.9	3.7	6.0
1F	20.5	17.0	12.0	11.9	7.5	7.5	4.0	5.5	3.4	3.7	3.1
2F	18.4	15.9	13.2	9.3	8.0	7.6	5.4	5.2	4.1	3.5	4.7
3F	21.2	16.3	12.6	13.0	7.2	7.3	5.4	4.3	5.9	3.3	4.5
4F	18.0	18.2	12.5	11.4	8.1	9.4	4.4	4.8	4.8	4.0	5.2

TABLE 5 - continued

Food conversion efficiency - group mean values (%)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group/ Sex	Week 12	Week 13	Week 1 to 13
1M	4.3	1.1	10.3
2M	5.3	0.5	10.8
3M	4.2	*	11.2
4M	4.1	0.4	10.7
1F	0.8	2.6	7.7
2F	0.4	4.7	7.7
3F	1.3	1.6	8.0
4F	0.9	1.2	7.9

* Bodyweight stasis or loss, food conversion efficiency not calculable

TABLE 6

Water consumption - group mean values (ml/animal/day)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Statistic	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11
1M	Mean	36	34	36	34	38	36	37	38	42	42	41
	SD	1.6	0.8	1.5	1.3	1.5	0.7	1.0	0.8	1.3	1.4	0.3
	N	2	2	2	2	2	2	2	2	2	2	2
2M	Mean	33	32	33	32	35	33	32	32	33	36	40
	SD	1.4	1.6	2.4	1.8	0.7	1.5	1.0	2.4	0.8	0.7	1.9
	N	2	2	2	2	2	2	2	2	2	2	2
3M	Mean	35	34	36	34	36	37	36	38	39	36	37
	SD	2.3	1.4	4.2	2.5	4.2	5.1	3.9	5.2	5.6	5.1	6.3
	N	2	2	2	2	2	2	2	2	2	2	2
4M	Mean	34	33	34	32	37	32	36	40	37	35	35
	SD	0.3	1.3	2.0	2.8	3.2	1.3	3.2	2.0	0.8	2.6	1.7
	N	2	2	2	2	2	2	2	2	2	2	2

TABLE 6 - continued

Water consumption - group mean values (ml/animal/day)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Statistic	Week 12	Week 13	Mean 1-13	As % of Control
1M	Mean	39	37	38	-
	SD	0.2	0.9		
	N	2	2		
2M	Mean	34	31	34	89
	SD	1.0	0.2		
	N	2	2		
3M	Mean	36	34	36	95
	SD	3.8	5.7		
	N	2	2		
4M	Mean	31	33	35	92
	SD	1.6	2.8		
	N	2	2		

TABLE 6 - continued

Water consumption - group mean values (ml/animal/day)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Statistic	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11
1F	Mean	24	25	26	24	28	27	28	25	31	29	28
	SD	0.8	2.4	2.1	1.7	1.6	1.3	1.7	1.8	1.6	1.7	3.7
	N	2	2	2	2	2	2	2	2	2	2	2
2F	Mean	23	25	26	25	27	26	28	27	29	27	26
	SD	0.2	2.5	1.8	1.2	2.3	2.8	0.0	2.3	2.9	2.3	2.3
	N	2	2	2	2	2	2	2	2	2	2	2
3F	Mean	21	23	23	23	25	23	25	25	25	24	22
	SD	0.5	0.9	0.8	0.7	1.3	1.7	1.5	2.1	0.5	1.3	0.9
	N	2	2	2	2	2	2	2	2	2	2	2
4F	Mean	25	25	28	26	33	27	31	35	30	29	29
	SD	1.2	0.3	0.5	1.2	1.6	1.3	1.5	5.3	1.0	1.9	0.1
	N	2	2	2	2	2	2	2	2	2	2	2

TABLE 6 - continued

Water consumption - group mean values (ml/animal/day)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Statistic	Week 12	Week 13	Mean 1-13	As % of Control
1F	Mean	24	26	27	-
	SD	1.8	0.2		
	N	2	2		
2F	Mean	26	25	26	96
	SD	3.2	2.2		
	N	2	2		
3F	Mean	23	22	23	85
	SD	1.6	1.7		
	N	2	2		
4F	Mean	25	27	28	104
	SD	1.7	0.0		
	N	2	2		

TABLE 7

Haematology - group mean values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex		Hct L/L	Hb g/dL	RBC $x10^{12}$/L	MCH pg	MCHC g/dL	MCV fL
1M	Mean	0.472	15.9	8.56	18.6	33.7	55.1
	SD	0.0127	0.36	0.315	0.58	0.47	1.09
	n	10	10	10	10	10	10
2M	Mean	0.467	15.7	8.54	18.4	33.6	54.7
	SD	0.0203	0.76	0.431	0.47	0.57	1.09
	n	10	10	10	10	10	10
3M	Mean	0.462	15.6	8.51	18.4	33.8	54.4
	SD	0.0114	0.29	0.203	0.36	0.47	1.19
	n	9	9	9	9	9	9
4M	Mean	0.450 b	15.2	8.29	18.3	33.7	54.3
	SD	0.0183	0.72	0.359	0.47	0.37	1.09
	n	10	10	10	10	10	10

Significant when compared with Group 1: a - $p<0.05$; b - $p<0.01$

TABLE 7 - continued

Haematology - group mean values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex		WBC $x10^9$/L	N $x10^9$/L	L $x10^9$/L	E $x10^9$/L	B $x10^9$/L	M $x10^9$/L	LUC $x10^9$/L
1M	Mean	12.61	1.48	10.64	0.13	0.06	0.26	0.04
	SD	2.233	0.513	2.177	0.032	0.024	0.095	0.013
	n	10	10	10	10	10	10	10
2M	Mean	13.52	1.88	11.08	0.16	0.08	0.29	0.04
	SD	2.935	1.227	1.985	0.055	0.024	0.112	0.028
	n	10	10	10	10	10	10	10
3M	Mean	11.01	1.66	8.85	0.12	0.06	0.29	0.03
	SD	1.775	0.654	1.418	0.050	0.016	0.068	0.011
	n	9	9	9	9	9	9	9
4M	Mean	11.68	1.50	9.68	0.14	0.06	0.26	0.03
	SD	2.407	0.646	2.262	0.063	0.024	0.064	0.016
	n	10	10	10	10	10	10	10

$p \geq 0.05$, no statistical significance

TABLE 7 - continued

Haematology - group mean values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex		Plt $x10^9$/L	PT sec	APTT sec
1M	Mean	1135	15.3	16.3
	SD	170.0	0.82	1.65
	n	10	10	10
2M	Mean	1138	15.8	15.4
	SD	230.9	0.67	1.70
	n	10	10	10
3M	Mean	1123	15.8	16.4
	SD	87.4	0.93	2.36
	n	9	9	9
4M	Mean	1123	15.9	17.7
	SD	116.7	0.84	2.98
	n	10	10	10

$p \geq 0.05$, no statistical significance

TABLE 7 - continued

Haematology - group mean values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex		Hct L/L	Hb g/dL	RBC $x10^{12}$/L	MCH pg	MCHC g/dL	MCV fL
1F	Mean	0.439	15.1	7.85	19.2	34.3	56.0
	SD	0.0147	0.65	0.363	0.52	0.54	1.50
	n	10	10	10	10	10	10
2F	Mean	0.450	15.3	7.94	19.3	34.0	56.6
	SD	0.0176	0.56	0.293	0.34	0.34	1.03
	n	10	10	10	10	10	10
3F	Mean	0.449	15.3	7.99	19.2	34.1	56.2
	SD	0.0179	0.49	0.243	0.38	0.55	0.90
	n	10	10	10	10	10	10
4F	Mean	0.441	15.1	7.82	19.3	34.2	56.5
	SD	0.0088	0.32	0.235	0.48	0.37	1.17
	n	10	10	10	10	10	10

$p \geq 0.05$, no statistical significance

TABLE 7 - continued

Haematology - group mean values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex		WBC $x10^9$/L	N $x10^9$/L	L $x10^9$/L	E $x10^9$/L	B $x10^9$/L	M $x10^9$/L	LUC $x10^9$/L
1F	Mean	8.47	0.77	7.39	0.10	0.03	0.15	0.02
	SD	2.351	0.359	2.066	0.028	0.014	0.036	0.011
	n	10	10	10	10	10	10	10
2F	Mean	9.52	1.49	7.61	0.12	0.03	0.24	0.03
	SD	3.232	1.258	2.187	0.049	0.020	0.191	0.014
	n	10	10	10	10	10	10	10
3F	Mean	8.20	0.64	7.29	0.09	0.02	0.12	0.03
	SD	2.093	0.320	1.869	0.017	0.010	0.042	0.015
	n	10	10	10	10	10	10	10
4F	Mean	7.26	0.94	6.06	0.09	0.02	0.13	0.02
	SD	3.429	0.826	2.779	0.041	0.021	0.084	0.018
	n	10	10	10	10	10	10	10

$p \geq 0.05$, no statistical significance

TABLE 7 - continued

Haematology - group mean values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex		Plt $x10^9$/L	PT sec	APTT sec
1F	Mean	1041	15.3	13.9
	SD	145.6	0.58	2.45
	n	10	9	9
2F	Mean	1072	15.5	14.3
	SD	99.0	0.41	2.50
	n	10	10	10
3F	Mean	1118	15.4	15.8
	SD	78.4	0.44	1.37
	n	10	10	10
4F	Mean	1083	15.3	15.8
	SD	72.9	0.43	1.62
	n	10	10	10

$p \geq 0.05$, no statistical significance

TABLE 8

Blood chemistry - group mean values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex		ALP u/L	ALT u/L	AST u/L	gGT u/L	Bili µmol/L	Urea mmol/L	Creat µmol/L	Gluc mmol/L	Chol mmol/L	Trig mmol/L
1M	Mean	92	42	68	0	2	5.44	34	6.48	1.71	0.88
	SD	13.6	5.6	7.1	0.5	0.4	0.845	5.0	0.917	0.282	0.300
	n	10	10	10	10	10	10	10	10	10	10
2M	Mean	89	42	67	1	2	5.37	33	7.32 a	1.79	1.02
	SD	19.3	6.3	12.4	0.5	0.7	0.914	3.0	1.074	0.299	0.472
	n	10	10	10	10	10	10	10	10	10	10
3M	Mean	94	42	66	0	2	5.71	34	7.52 a	1.60	0.82
	SD	20.6	7.1	8.7	0.4	0.7	0.784	2.6	0.508	0.159	0.257
	n	9	9	9	9	9	9	9	9	9	9
4M	Mean	84	42	67	1	2	5.80	34	7.38 a	1.72	0.78
	SD	13.6	5.3	8.8	0.7	0.7	0.981	4.8	0.793	0.368	0.245
	n	10	10	10	10	10	10	10	10	10	10

Significant when compared with Group 1: a - $p<0.05$; b - $p<0.01$

TABLE 8 - continued

Blood chemistry - group mean values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex		Na mmol/L	K mmol/L	Cl mmol/L	Ca mmol/L	Phos mmol/L	Total Prot g/L	Alb g/L	A/G Ratio
1M	Mean	140	5.6	100	2.49	2.07	70	39	1.26
	SD	2.3	0.65	2.8	0.075	0.158	3.6	1.2	0.133
	n	10	10	10	10	10	10	10	10
2M	Mean	139	5.6	99	2.53	2.00	68	38	1.26
	SD	1.9	0.67	1.3	0.089	0.194	4.4	2.1	0.093
	n	10	10	10	10	10	10	10	10
3M	Mean	140	5.2	99	2.50	1.94	69	37	1.19
	SD	1.6	0.74	1.3	0.053	0.184	3.1	0.7	0.115
	n	9	9	9	9	9	9	9	9
4M	Mean	140	5.5	100	2.45	1.91	66	36 B	1.24
	SD	1.8	0.64	1.9	0.066	0.129	2.9	1.3	0.144
	n	10	10	10	10	10	10	10	10

Significant when compared with Group 1: A - $p<0.05$; B - $p<0.01$

TABLE 8 - continued

Blood chemistry - group mean values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex		ALP u/L	ALT u/L	AST u/L	gGT u/L	Bili µmol/L	Urea mmol/L	Creat µmol/L	Gluc mmol/L	Chol mmol/L	Trig mmol/L
1F	Mean	46	34	70	1	2	5.64	39	6.01	2.43	0.75
	SD	10.1	8.5	28.3	0.5	0.5	0.830	3.3	0.607	0.367	0.431
	n	10	10	10	10	10	10	10	10	10	10
2F	Mean	47	48	78	1	2	6.19	45 b	6.30	2.33	0.96
	SD	11.1	31.3	21.5	0.7	0.6	0.637	3.0	0.934	0.322	0.268
	n	10	10	10	10	10	10	10	10	10	10
3F	Mean	48	35	76	1	2	6.21	47 b	6.62	2.14	0.86
	SD	7.7	6.3	11.2	0.8	0.7	0.756	3.6	1.114	0.561	0.320
	n	10	10	10	10	10	10	10	10	10	10
4F	Mean	46	34	69	1	2	6.51	48 b	6.93	2.22	1.01
	SD	15.2	5.7	7.5	0.4	0.6	1.090	6.2	0.883	0.461	0.589
	n	10	10	10	10	10	10	10	10	10	10

Significant when compared with Group 1: a - $p<0.05$; b - $p<0.01$

TABLE 8 - continued

Blood chemistry - group mean values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex		Na mmol/L	K mmol/L	Cl mmol/L	Ca mmol/L	Phos mmol/L	Total Prot g/L	Alb g/L	A/G Ratio
1F	Mean	139	4.3	99	2.61	1.80	74	43	1.39
	SD	1.3	0.59	2.2	0.054	0.240	5.3	4.0	0.189
	n	10	10	10	10	10	10	10	10
2F	Mean	139	4.9	100	2.59	1.72	73	42	1.36
	SD	2.2	0.69	2.1	0.078	0.197	4.7	2.8	0.167
	n	10	10	10	10	10	10	10	10
3F	Mean	140	4.9	99	2.53 a	1.68	71	41	1.35
	SD	1.5	0.78	1.8	0.067	0.189	4.0	1.9	0.114
	n	10	10	10	10	10	10	10	10
4F	Mean	140	4.9	100	2.53 a	1.54	74	42	1.30
	SD	1.5	0.41	1.1	0.070	0.293	3.4	2.4	0.105
	n	10	10	10	10	10	10	10	10

Significant when compared with Group 1: a - $p<0.05$; b - $p<0.01$

TABLE 9

Organ weights - group mean absolute values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Terminal bodyweight	Brain	Thymus	Heart	Liver	Spleen	Adrenals	Kidneys	Testes
1M	Mean	515	2.18	0.242	1.511	17.10	0.776	0.048	2.99	3.57
	SD	66.4	0.088	0.0837	0.1714	3.951	0.1017	0.0068	0.500	0.186
	N	10	10	10	10	10	10	10	10	10
2M	Mean	527	2.10	0.242	1.509	17.69	0.810	0.050	3.05	3.54
	SD	56.3	0.081	0.0563	0.1731	2.066	0.1130	0.0142	0.276	0.274
	N	10	10	10	10	10	10	10	10	10
3M	Mean	519	2.07	0.206	1.527	16.86	0.757	0.049	3.12	3.59
	SD	58.9	0.109	0.0655	0.2013	2.659	0.1346	0.0114	0.354	0.222
	N	9	9	9	9	9	9	9	9	9
4M	Mean	505	2.12	0.237	1.514	16.20	0.772	0.046	2.95	3.66
	SD	59.3	0.105	0.0608	0.1165	2.999	0.1778	0.0045	0.388	0.245
	N	10	10	10	10	10	10	10	10	10

TABLE 9 - continued

Organ weights - group mean absolute values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Epididymides
1M	Mean	1.265
	SD	0.0813
	N	10
2M	Mean	1.215
	SD	0.1292
	N	10
3M	Mean	1.240
	SD	0.0738
	N	9
4M	Mean	1.248
	SD	0.0956
	N	10

TABLE 9 - continued

Organ weights - group mean absolute values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Terminal bodyweight	Brain	Thymus	Heart	Liver	Spleen	Adrenals	Kidneys	Ovaries
1F	Mean	305	2.02	0.233	1.026	10.50	0.572	0.066	1.99	0.087
	SD	21.1	0.070	0.0569	0.0652	1.022	0.0943	0.0070	0.201	0.0091
	N	10	10	10	10	10	10	10	10	10
2F	Mean	302	1.99	0.274	1.013	10.09	0.547	0.063	1.86	0.085
	SD	21.3	0.067	0.0555	0.0637	0.891	0.0831	0.0077	0.193	0.0104
	N	10	10	10	10	10	10	10	10	10
3F	Mean	301	2.00	0.284	0.969	9.65	0.565	0.059	1.89	0.087
	SD	26.9	0.101	0.0664	0.0876	1.094	0.0875	0.0097	0.234	0.0159
	N	10	10	10	10	10	10	10	10	10
4F	Mean	315	2.01	0.231	1.051	10.25	0.550	0.063	1.93	0.090
	SD	35.2	0.080	0.0764	0.1084	1.123	0.0848	0.0066	0.224	0.0142
	N	10	10	10	10	10	10	10	10	10

TABLE 9 - continued

Organ weights - group mean absolute values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Uterus + Cervix
1F	Mean	0.720
	SD	0.1856
	N	10
2F	Mean	0.678
	SD	0.1657
	N	10
3F	Mean	0.685
	SD	0.1993
	N	10
4F	Mean	0.650
	SD	0.1910
	N	10

TABLE 9 - continued

Organ weights - group mean adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Terminal bodyweight	Brain	Thymus	Heart	Liver	Spleen	Adrenals	Kidneys	Testes
1M	Mean	515	2.18	0.24	1.513	17.15	0.777	0.048	2.99	3.569
	SD	66.4	0.088	0.084	0.1714	3.951	0.1017	0.0068	0.500	0.1862
	N	10	10	10	10	10	10	10	10	10
2M	Mean	527	2.10	0.23	1.486	17.21	0.796	0.050	2.99	3.538
	SD	56.3	0.081	0.056	0.1731	2.066	0.1130	0.0142	0.276	0.2742
	N	10	10	10	10	10	10	10	10	10
3M	Mean	519	2.07	0.20	1.522	16.76	0.754	0.049	3.11	3.586
	SD	58.9	0.109	0.065	0.2013	2.659	0.1346	0.0114	0.354	0.2222
	N	9	9	9	9	9	9	9	9	9
4M	Mean	505	2.13	0.25	1.539	16.72	0.786	0.046	3.00	3.658
	SD	59.3	0.105	0.061	0.1165	2.999	0.1778	0.0045	0.388	0.2449
	N	10	10	10	10	10	10	10	10	10

$p \geq 0.05$, no statistical significance

TABLE 9 - continued

Organ weights - group mean adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Epididymides
1M	Mean	1.265
	SD	0.0813
	N	10
2M	Mean	1.215
	SD	0.1292
	N	10
3M	Mean	1.240
	SD	0.0738
	N	9
4M	Mean	1.248
	SD	0.0956
	N	10

$p \geq 0.05$, no statistical significance

TABLE 9 - continued

Organ weights - group mean adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Terminal bodyweight	Brain	Thymus	Heart	Liver	Spleen	Adrenals	Kidneys	Ovaries
1F	Mean	305	2.02	0.233	1.028	10.53	0.574	0.066	1.99	0.087
	SD	21.1	0.070	0.0569	0.0652	1.022	0.0943	0.0070	0.201	0.0091
	N	10	10	10	10	10	10	10	10	10
2F	Mean	302	1.99	0.277	1.020	10.18	0.553	0.063	1.88	0.085
	SD	21.3	0.067	0.0555	0.0637	0.891	0.0831	0.0077	0.193	0.0104
	N	10	10	10	10	10	10	10	10	10
3F	Mean	301	2.00	0.288	0.978	9.78	0.573	0.059	1.91	0.087
	SD	26.9	0.101	0.0664	0.0876	1.094	0.0875	0.0097	0.234	0.0159
	N	10	10	10	10	10	10	10	10	10
4F	Mean	315	2.01	0.222	1.032	9.99	0.534	0.063	1.89	0.090
	SD	35.2	0.080	0.0764	0.1084	1.123	0.0848	0.0066	0.224	0.0142
	N	10	10	10	10	10	10	10	10	10

$p \geq 0.05$, no statistical significance

TABLE 9 - continued

Organ weights - group mean adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex		Uterus + Cervix
1F	Mean	0.720
	SD	0.1856
	N	10
2F	Mean	0.678
	SD	0.1657
	N	10
3F	Mean	0.685
	SD	0.1993
	N	10
4F	Mean	0.650
	SD	0.1910
	N	10

$p \geq 0.05$, no statistical significance

TABLE 10

Macropathology - group distribution of findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Observation	Group/Sex	1M	2M	3M	4M	1F	2F	3F	4F
	Number examined	10	10	9	10	10	10	10	10
Adrenals									
Pale Area(s)		0	0	0	0	1	0	0	0
Kidneys									
Enlarged		0	0	1	0	0	0	0	0
Pelvic dilatation		0	0	1	0	2	1	0	0
Depression(s)		0	0	0	1	0	2	0	0
Pale area(s)		0	0	1	1	0	0	0	0
Liver									
Mass(es)		0	0	1	0	0	0	0	0
Pale area(s)		0	1	0	0	0	0	0	0
Swollen		1	0	0	0	0	0	0	0
Ruptured		0	0	1	0	0	0	0	0
Lungs + Bronchi									
Pale area(s)		0	0	0	0	1	0	0	2
LN Mandibular									
Enlarged		2	2	3	2	1	5	1	1
Salivary Glands									
Oedematous		1	0	0	0	0	0	0	0

$p \geq 0.05$, no statistical significance

TABLE 10 - continued

Macropathology - group distribution of findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Observation	Group/Sex	1M	2M	3M	4M	1F	2F	3F	4F
	Number examined	10	10	9	10	10	10	10	10
Skin									
Scab(s)		0	0	0	0	2	0	0	0
Stomach									
Antrum white nodule(s)		0	0	0	0	0	1	1	0
Thymus									
Dark area(s)		0	1	3	1	0	0	0	0
General Comments									
Moderate hairloss		1	0	0	0	5	3	2	1
Tail partially absent		0	1	0	0	0	0	0	0
Fur stained		0	0	2	0	0	0	0	0
Adipose tissue									
Strangulated Nodule(s)		0	1	0	0	0	0	0	0
Abdomen									
Contained fluid		0	0	1	0	0	0	0	0
Uterus									
Fluid distension		0	0	0	0	4	3	2	3

$p \geq 0.05$, no statistical significance

TABLE 11

Histopathology - group distribution of findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Observation	Group/Sex	1M	2M	3M	4M	1F	2F	3F	4F
Trachea	Number examined:	10	0	0	10	10	0	0	10
Subepithelial Lymphocytic Infiltration		0	0	0	0	0	0	0	1
Lungs + Bronchi	Number examined:	10	0	0	10	10	0	0	10
Arterial Mural Mineralisation		3	0	0	7	1	0	0	2
Alveolar Macrophages		2	0	0	2	1	0	0	2
Alveolitis		0	0	0	1	1	0	0	0
Alveolar Epithelial Hyperplasia		0	0	0	1	0	0	0	0
Cholesterol Cleft Granuloma(ta)		0	0	0	0	1	0	0	0
Liver	Number examined:	10	1	2	10	10	0	0	10
Arteritis/Periarteritis		0	0	0	1	0	0	0	0
Inflammation, Portal		0	0	0	0	1	0	0	0
Subcapsular Necrosis With Inflammation		0	1	0	0	0	0	0	0
Kidneys	Number examined:	10	0	1	10	10	2	0	10
Interstitial Inflammatory Cells		0	0	1	0	1	0	0	1
Cortical Tubules with Hyaline Droplets		4	0	0	6	0	0	0	0
Cortical Tubular Basophilia		3	0	0	4	0	0	0	0
Transitional Epithelial Inflammatory Cells		0	0	0	1	0	0	0	0
Transitional Epithelial Hyperplasia		0	0	0	1	0	0	0	0
Hydronephrosis		0	0	1	0	2	1	0	0
Medullary Tubular Dilatation		0	0	0	0	1	0	0	0

$p \geq 0.05$, no statistical significance

TABLE 11 - continued

Histopathology - group distribution of findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Observation	Group/Sex	1M	2M	3M	4M	1F	2F	3F	4F
Kidneys	Number examined:	10	0	1	10	10	2	0	10
Mineralisation, Corticomedullary		0	0	0	0	1	0	0	0
Urinary Bladder	Number examined:	10	0	0	10	10	0	0	10
Luminal Dilatation		4	0	0	0	0	0	0	0
Heart	Number examined:	10	0	0	10	10	0	0	10
Myocardial Inflammatory Cells		3	0	0	3	2	0	0	4
Epicardial Inflammatory Cells		0	0	0	1	0	0	0	0
Myocardial Fibrosis		0	0	0	2	0	0	0	0
Endocardial Inflammatory Cells		0	0	0	1	0	0	0	0
Spleen	Number examined:	10	0	0	10	10	0	0	10
Extramedullary Haemopoiesis		5	0	0	4	3	0	0	2
Pancreas	Number examined:	10	0	0	10	10	0	0	10
Perivascular Inflammatory Cells		3	0	0	4	2	0	0	5
Basophilic Acinar Cells, Focal		0	0	0	0	0	0	0	1
Periductal Inflammation / Fibrosis		0	0	0	1	0	0	0	1
Acinar Replacement by Adipose Tissue		0	0	0	1	0	0	0	0
Eosinophil Infiltration		0	0	0	1	0	0	0	0
Oedema		0	0	0	1	0	0	0	0
Acinar Atrophy, Focal		0	0	0	1	0	0	0	0

$p \geq 0.05$, no statistical significance

TABLE 11 - continued

Histopathology - group distribution of findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Observation	Group/Sex	1M	2M	3M	4M	1F	2F	3F	4F
LN Mesenteric	Number examined:	10	0	0	10	10	0	0	10
Sinus Erythrocytosis/Erythrophagocytosis		1	0	0	1	1	0	0	0
Mastocytosis		1	0	0	0	0	0	0	0
Thymus	Number examined:	9	1	3	10	10	0	0	10
Haemorrhage		0	0	3	0	0	0	0	0
Thyroids	Number examined:	10	0	0	10	10	0	0	10
Follicular Cell Hypertrophy		5	0	0	3	0	0	0	0
Ectopic Thymic Tissue		1	0	0	0	0	0	0	0
Adrenals	Number examined:	10	10	9	10	10	0	0	10
Cortical Vacuolation		2	2	4	5	0	0	0	0
Cortical Hypertrophy with Vacuolation, Focal		0	0	0	0	1	0	0	0
Pituitary	Number examined:	10	0	0	10	10	0	0	10
Perivascular Inflammatory Cells		0	0	0	0	0	0	0	1
Vacuolated Basophilic Cells		4	0	0	3	0	0	0	0
Developmental Cyst(s)		0	0	0	1	0	0	0	0
LN Mandibular	Number examined:	10	2	3	10	10	5	1	10
Reactive Histiocytosis		0	0	0	0	0	0	0	1
Plasmacytosis		6	2	3	6	3	5	1	0

$p \geq 0.05$, no statistical significance

TABLE 11 - continued

Histopathology - group distribution of findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Observation	Group/Sex	1M	2M	3M	4M	1F	2F	3F	4F
LN Mandibular	Number examined:	10	2	3	10	10	5	1	10
Mastocytosis		1	0	0	1	0	0	0	0
Sinus Erythrocytosis/Erythrophagocytosis		0	0	0	2	0	0	0	0
Salivary Glands	Number examined:	10	0	0	10	10	0	0	10
Inflammation		0	0	0	0	0	0	0	2
Stomach	Number examined:	10	0	0	10	10	1	1	10
Mucosal Inflammation - Glandular Region		0	0	0	1	0	0	0	0
Ectopic Nonglandular Epithelium in Glandular Region, Focal		0	0	0	0	0	0	1	0
Skin	Number examined:	10	0	0	10	10	0	0	10
Scab(s)		0	0	0	0	1	0	0	0
Epidermal Ulceration		0	0	0	0	1	0	0	0
Epidermal Hyperplasia		0	0	0	0	1	0	0	0
Femur inc. Joint	Number examined:	10	0	0	10	10	0	0	10
Marrow - Fat Replacement		7	0	0	7	2	0	0	0
Spinal C. Lumb.	Number examined:	10	0	0	10	10	0	0	10
Degenerate Fibres		0	0	0	0	0	0	0	1

$p \geq 0.05$, no statistical significance

TABLE 11 - continued

Histopathology - group distribution of findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Observation	Group/Sex	1M	2M	3M	4M	1F	2F	3F	4F
Prostate	Number examined:	10	0	0	10	0	0	0	0
Inflammation		2	0	0	3	0	0	0	0
Uterus	Number examined:	0	0	0	0	10	3	2	10
Luminal Dilatation		0	0	0	0	4	3	2	3
Epididymides	Number examined:	10	0	0	10	0	0	0	0
Interstitial Inflammatory Cells		4	0	0	2	0	0	0	0
Adipose tissue	Number examined:	0	1	0	0	0	0	0	0
Fat Necrosis		0	1	0	0	0	0	0	0

$p \geq 0.05$, no statistical significance

TABLE 11 - continued

Historical Histopathology Data -
Incidence of cortical vacuolation of the Adrenals in Control male CD rats from recent toxicity studies performed at Huntingdon Life Sciences

Start date	Jun-04	Nov-04	Nov-04	Jan-05	Feb-05	Oct-05	Dec-05			
Route of administration.	og	og	og	ih	og	ih	og			
Supplier	CRUK	CRUK	CRUK	CRUK	CRUK	CRUK	CRUK			
Number of animals per cage	5	3	3	4	5	5	5			
Study duration (weeks)	13	13	13	13	13	13	13			
								total	**range of percentages***	
									min	**max**
Number of animals	10	12	12	12	10	5	10	71		
Adrenals										
Number examined	10	12	12	12	10	5	10	71		
Cortical Vacuolation										
minimal	0	0	0	4	0	2	4	10		
slight	3	0	2	1	0	0	0	6		
moderate	0	0	0	0	0	0	0	0		
total	3	0	2	5	0	2	4	16		
percentage*	30.0%	0.0%	16.7%	41.7%	0.0%	40.0%	40.0%	22.54%	0.00%	41.67%

CRUK - Charles River (UK) Ltd
og - Oral gavage
ih - Inhalation

Mortality, macropathology and histopathology - individual findings for animals killed during the treatment period

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 10 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Killed for welfare reasons
Date of Death: 17-Dec-05 Day of Death: 20 Week of Death: 3 Subgroup: 1 Terminal Bodyweight: 341.8 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Adrenals Cortical Hypertrophy, Moderate Hypertrophy was mainly in zona fasciculata
Brain Swollen	Brain Meningitis, Marked
	Liver Hepatocyte Vacuolation, Periportal, Slight
	Pancreas Acinar Cell Apoptosis, Minimal
	Spinal C. Cerv. Meningitis, Minimal
	Spinal C. Thor. Meningitis, Minimal
	Spinal C. Lumb. Meningitis, Minimal

Mortality, macropathology and histopathology - individual findings for animals killed during the treatment period

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 10 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Killed for welfare reasons
Date of Death: 17-Dec-05 Day of Death: 20 Week of Death: 3 Subgroup: 1 Terminal Bodyweight: 341.8 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
continued from previous page...	Salivary Glands Acinar Cell Hypertrophy, Moderate Acinar cell hypertrophy-submandibular/sublingual glands Thymus Apoptosis - Cortex, Minimal Urinary Bladder Luminal Dilatation, Present
General Comments Abnormal contents GI tract, Ileum, jejunum and duodenum contain yellow, viscous fluid. Fur stained, Muzzle, Brown Fur stained, Orbit(s), Brown	General Comments No significant lesions

Signs - individual observations in association with dosing

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Sign	Pre-dose	Days On return to cage	On completion of dosing each group	1 to 2 hours after dosing	As late as possible in working day
2M	38	Behaviour					
		Salivation, Slight			1		
2F	43	Behaviour					
		Vocalising		5			
4F	71	Behaviour					
		Vocalising		5	5	5	

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
1M	11	T	92		No abnormalities detected	
	12	T	92		No abnormalities detected	
	13	T	92		No abnormalities detected	
	14	T	92		No abnormalities detected	
	15	T	92	Skin Colour	Reddening, Nose	6
	16	T	92	Coat	Hairloss, Forelimbs	9-13
				Skin	Encrustation, Upper Dorsal Thorax (Left)	3
	17	T	92	Coat	Hairloss, Head	11-14
				Skin	Encrustation, Head	7-10
				Staining	Abnormal Colour, Upper Dorsal Thorax	6-10
					Abnormal Colour, Brown, Upper Dorsal Thorax	12
	18	T	92		No abnormalities detected	
	19	T	92	Coat	Hairloss, Dorsal Body Surface	12
	20	T	92	Coat	Hairloss, Head	9-14

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
2M	31	T	92		No abnormalities detected	
	32	T	92		No abnormalities detected	
	33	T	92		No abnormalities detected	
	34	T	92		No abnormalities detected	
	35	T	92		No abnormalities detected	
	36	T	92		No abnormalities detected	
	37	T	92	Staining	Abnormal Colour, Brown, Eye (Right)	1
	38	T	92		No abnormalities detected	
	39	T	92		No abnormalities detected	
	40	T	92	Build Deformity	Partially absent appendage, Tail	4-14
				Skin	Encrustation, Tail	12-13

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
3M	1	T	92		No abnormalities detected	
	2	T	92		No abnormalities detected	
	3	T	92		No abnormalities detected	
	4	T	92		No abnormalities detected	
	5	T	92	Staining	Abnormal Colour, Brown, Orbital (Left)	1
					Abnormal Colour, Brown, Orbit (Right)	5
					Abnormal Colour, Orbital	4
	6	T	92	Coat	Hairloss, Forelimbs	5
					Hairloss, Dorsal Body Surface	12-13
				Skin	Encrustation, Upper Dorsal Thorax	12-13
	7	T	92	Staining	Abnormal Colour, Brown, Upper Dorsal Thorax	12-14
	8	T	92	Staining	Abnormal Colour, Brown, Upper Dorsal Thorax	12-14
					Abnormal Colour, Brown, Head	12-14
	9	T	92	Build Deformity	Swollen Area, Muzzle	3-4

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
3M	10	W	20	Behaviour	Underactive	3
				Eyelids	Partially Closed, Upper Eyelid (Left)	3
					Partially Closed, Lower Eyelid (Left)	3
					Partially Closed, Upper Eyelid (Right)	3
					Partially Closed, Lower Eyelid (Right)	3
				Muscle Reaction	Tremor	3
					Abnormal Gait	3
				Staining	Abnormal Colour, Brown, Eyes	3
					Abnormal Colour, Brown, Nose	3

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
4M	21	T	92		No abnormalities detected	
	22	T	92		No abnormalities detected	
	23	T	92		No abnormalities detected	
	24	T	92		No abnormalities detected	
	25	T	92	Skin	Encrustation, Upper Dorsal Thorax (Left)	3
	26	T	92		No abnormalities detected	
	27	T	92		No abnormalities detected	
	28	T	92		No abnormalities detected	
	29	T	92		No abnormalities detected	
	30	T	92		No abnormalities detected	

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
1F	51	T	93	Coat	Hairloss, Head	12-14
				Skin	Encrustation, Head	12, 14
	52	T	93		No abnormalities detected	
	53	T	93	Coat	Hairloss, Head	13-14
				Skin	Encrustation, Head	14
	54	T	93		No abnormalities detected	
	55	T	93	Coat	Hairloss, Head	12-14
				Skin	Encrustation, Head	12
				Staining	Abnormal Colour, Brown, Upper Dorsal Thorax	5
	56	T	93	Coat	Hairloss, Head	5, 14
	57	T	93		No abnormalities detected	
	58	T	93	Coat	Hairloss, Head	12-14
	59	T	93	Coat	Hairloss, Head	9-14
				Skin	Encrustation, Head	12-14

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
1F	60	T	93	Coat	Hairloss, Head	12-13

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
2F	41	T	93		No abnormalities detected	
	42	T	93	Coat	Hairloss, Head	11-14
				Skin	Encrustation, Head	12-13
	43	T	93	Behaviour	Vocalisation	1-14
	44	T	93	Behaviour	Vocalisation	1
				Coat	Hairloss, Head	12-14
	45	T	93	Skin Colour	Reddening, Pinna	3
	46	T	93	Coat	Hairloss, Head	9-12
					Hairloss, Ventral Body Surface	11-13
				Skin	Encrustation, Head	12-13
	47	T	93	Coat	Hairloss, Head	5-14
	48	T	93	Behaviour	Vocalisation	5
				Coat	Hairloss, Head	5-14
				Skin	Encrustation, Head	5-6

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
2F	49	T	93	Coat	Hairloss, Head	6-7
					Hairloss, Dorsal Body Surface	3-4, 12
				Skin	Encrustation, Upper Dorsal Thorax (Left)	3
					Encrustation, Muzzle	9
	50	T	93	Coat	Hairloss, Head	5-9
					Hairloss, Ventral Body Surface	12-13

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
3F	61	T	93	Staining	Abnormal Colour, Brown, Upper Dorsal Thorax	12-13
	62	T	93		No abnormalities detected	
	63	T	93		No abnormalities detected	
	64	T	93		No abnormalities detected	
	65	T	93		No abnormalities detected	
	66	T	93	Coat	Hairloss, Head	7-11
	67	T	93	Behaviour	Vocalisation	1-2
				Coat	Hairloss, Head	7-14
				Skin	Encrustation, Head	11-13
	68	T	93		No abnormalities detected	
	69	T	93	Coat	Hairloss, Head	9-14
	70	T	93	Coat	Hairloss, Head	9-14

Detailed physical examination and arena observations - individual observations

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal	Death Code	Death Day	Category	Observation	Week(s)
4F	71	T	93	Behaviour	Vocalisation	1
	72	T	93		No abnormalities detected	
	73	T	93		No abnormalities detected	
	74	T	93		No abnormalities detected	
	75	T	93		No abnormalities detected	
	76	T	93		No abnormalities detected	
	77	T	93	Behaviour	Vocalisation	1-2
	78	T	93	Coat	Hairloss, Forelimbs	12-14
	79	T	93	Behaviour	Vocalisation	2
	80	T	93		No abnormalities detected	

APPENDIX 3

Sensory reactivity observations and grip strength - individual findings during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group/sex:		1M										2M									
Animal number:		11	12	13	14	15	16	17	18	19	20	31	32	33	34	35	36	37	38	39	40
Parameter (Grade range)																					
Approach response (1-3)		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Touch response (1-3)		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Auditory startle reflex (1-4)		2	3	3	3	3	2	3	3	3	2	3	3	3	3	3	3	3	3	3	3
Tail pinch response (1-4)		2	3	2	2	2	2	2	3	2	2	3	3	3	3	3	3	3	3	2	3
Forelimb grip strength (kg)	Trial 1	1.35	1.46	1.38	1.39	1.42	1.48	1.35	1.25	1.60	1.68	1.06	1.28	0.88	1.10	1.02	1.15	1.20	1.18	1.37	1.51
	Trial 2	1.24	1.64	1.46	1.17	1.70	1.28	1.32	1.56	1.57	1.44	1.02	1.22	0.88	1.21	1.14	1.00	1.26	1.32	1.18	1.34
	Trial 3	1.23	1.52	1.48	1.48	1.49	1.44	1.31	1.32	1.36	1.50	1.09	1.14	0.96	1.27	1.12	1.17	1.22	1.17	1.35	1.54
	Mean	1.27	1.54	1.44	1.35	1.54	1.40	1.33	1.38	1.51	1.54	1.06	1.21	0.91	1.19	1.09	1.11	1.23	1.22	1.30	1.46
Hindlimb grip strength (kg)	Trial 1	0.68	0.68	0.46	0.59	0.51	0.60	0.48	0.55	0.70	0.74	0.39	0.59	0.48	0.53	0.71	0.76	0.85	0.67	0.68	0.74
	Trial 2	0.78	0.62	0.54	0.51	0.46	0.65	0.52	0.52	0.64	0.72	0.64	0.55	0.45	0.56	0.63	0.59	0.71	0.61	0.55	0.54
	Trial 3	0.78	0.70	0.49	0.60	0.48	0.66	0.51	0.49	0.65	0.65	0.56	0.62	0.38	0.43	0.75	0.73	0.68	0.71	0.60	0.47
	Mean	0.75	0.67	0.50	0.57	0.48	0.64	0.50	0.52	0.66	0.70	0.53	0.59	0.44	0.51	0.70	0.69	0.75	0.66	0.61	0.58
Additional observations		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Forelimb grip strength (kg)	Mean	1.43	1.18
	SD	0.10	0.15
	N	10	10
	Method†	2.7	
	P	0.000	0.000
Hindlimb grip strength (kg)	Mean	0.60	0.61
	SD	0.10	0.10
	N	10	10
	Method†	2.7	
	P	0.991	

† See Statistical Analysis Section in Experimental Procedure for explanation

APPENDIX 3 - continued

Sensory reactivity observations and grip strength - individual findings during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group/sex:		3M									4M									
Animal number:		1	2	3	4	5	6	7	8	9	21	22	23	24	25	26	27	28	29	30
Parameter (Grade range)																				
Approach response (1-3)		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Touch response (1-3)		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Auditory startle reflex (1-4)		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Tail pinch response (1-4)		3	3	3	2	3	3	4	3	2	3	3	4	3	4	3	3	3	2	3
Forelimb grip strength (kg)	Trial 1	1.27	1.28	1.39	1.34	1.36	1.16	1.67	1.16	1.20	1.50	1.22	1.10	1.04	1.35	1.14	1.27	1.20	1.17	1.33
	Trial 2	1.19	1.24	1.21	1.38	1.14	1.28	1.38	1.18	1.09	1.45	1.26	0.93	1.34	1.19	1.08	1.04	1.21	1.30	1.21
	Trial 3	1.30	1.06	1.05	1.19	1.22	1.33	1.27	1.29	1.33	1.39	1.32	1.08	0.93	1.24	1.14	1.17	1.19	1.29	1.30
	Mean	1.25	1.19	1.22	1.30	1.24	1.26	1.44	1.21	1.21	1.45	1.27	1.04	1.10	1.26	1.12	1.16	1.20	1.25	1.28
Hindlimb grip strength (kg)	Trial 1	0.50	0.72	0.58	0.54	0.41	0.58	0.63	0.63	0.57	0.66	0.43	0.67	0.57	0.59	0.49	0.50	0.64	0.73	0.53
	Trial 2	0.50	0.69	0.65	0.61	0.51	0.58	0.64	0.70	0.62	0.63	0.58	0.61	0.66	0.69	0.42	0.70	0.57	0.69	0.60
	Trial 3	0.54	0.50	0.58	0.82	0.47	0.66	0.64	0.76	0.67	0.64	0.46	0.64	0.65	0.49	0.48	0.62	0.57	0.72	0.64
	Mean	0.51	0.64	0.60	0.66	0.46	0.61	0.64	0.70	0.62	0.64	0.49	0.64	0.63	0.59	0.46	0.61	0.59	0.71	0.59
Additional observations		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Forelimb grip strength (kg)	Mean	1.26	1.21
	SD	0.08	0.12
	N	9	10
	P	0.002	0.000
Hindlimb grip strength (kg)	Mean	0.60	0.60
	SD	0.07	0.07
	N	9	10
	P		

APPENDIX 3 - continued

Sensory reactivity observations and grip strength - individual findings during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group/sex:		1F										2F									
Animal number:		51	52	53	54	55	56	57	58	59	60	41	42	43	44	45	46	47	48	49	50
Parameter (Grade range)																					
Approach response (1-3)		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Touch response (1-3)		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Auditory startle reflex (1-4)		3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Tail pinch response (1-4)		3	2	2	3	2	3	2	3	2	2	3	2	2	3	3	3	2	3	3	3
Forelimb grip strength (kg)	Trial 1	0.94	1.21	1.25	1.14	1.27	0.85	1.09	0.96	0.96	1.07	1.16	0.98	1.00	1.20	1.04	0.88	0.91	1.31	1.22	1.14
	Trial 2	1.21	1.30	1.14	1.06	1.20	0.89	1.17	1.26	1.00	1.17	1.13	1.36	1.15	1.34	1.18	0.77	0.98	1.35	1.25	1.02
	Trial 3	1.18	1.32	1.32	1.08	1.28	0.97	1.04	0.98	1.07	1.04	1.33	1.20	0.96	1.17	1.05	0.76	0.70	1.05	1.32	1.05
	Mean	1.11	1.28	1.24	1.09	1.25	0.90	1.10	1.07	1.01	1.09	1.21	1.18	1.04	1.24	1.09	0.80	0.86	1.24	1.26	1.07
Hindlimb grip strength (kg)	Trial 1	0.54	0.59	0.54	0.56	0.54	0.31	0.62	0.45	0.40	0.32	0.42	0.32	0.54	0.35	0.45	0.63	0.46	0.42	0.51	0.49
	Trial 2	0.41	0.40	0.60	0.47	0.55	0.40	0.47	0.62	0.53	0.32	0.49	0.43	0.40	0.35	0.52	0.62	0.44	0.45	0.40	0.40
	Trial 3	0.43	0.59	0.50	0.49	0.57	0.41	0.51	0.57	0.41	0.38	0.44	0.41	0.42	0.33	0.47	0.59	0.52	0.52	0.40	0.41
	Mean	0.46	0.53	0.55	0.51	0.55	0.37	0.53	0.55	0.45	0.34	0.45	0.39	0.45	0.34	0.48	0.61	0.47	0.46	0.44	0.43
Additional observations		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Forelimb grip strength (kg)	Mean	1.11	1.10
	SD	0.12	0.16
	N	10	10
	Method†	2.7	
	P	0.588	
Hindlimb grip strength (kg)	Mean	0.48	0.45
	SD	0.08	0.07
	N	10	10
	Method†	2.7	
	P	0.535	

† See Statistical Analysis Section in Experimental Procedure for explanation

APPENDIX 3 - continued

Sensory reactivity observations and grip strength - individual findings during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group/sex:		3F										4F									
Animal number:		61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
Parameter (Grade range)																					
Approach response (1-3)		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Touch response (1-3)		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Auditory startle reflex (1-4)		3	2	3	2	3	3	3	3	3	3	2	3	3	3	3	3	3	3	3	3
Tail pinch response (1-4)		3	2	3	3	3	3	3	3	3	2	3	3	3	2	3	2	3	3	3	3
Forelimb grip strength (kg)	Trial 1	1.04	1.11	1.22	1.16	0.93	1.16	0.98	0.88	1.10	1.10	0.90	1.16	0.98	1.13	1.02	1.06	1.35	0.90	1.06	0.98
	Trial 2	0.98	0.93	1.03	1.38	1.01	1.06	1.01	0.82	0.93	1.31	0.96	1.04	0.95	1.12	1.23	1.08	1.24	0.92	1.08	0.96
	Trial 3	1.07	1.09	1.52	1.08	0.93	1.16	0.89	1.02	0.92	1.16	0.86	1.06	0.99	1.05	1.27	1.08	1.16	0.69	1.04	0.81
	Mean	1.03	1.04	1.26	1.21	0.96	1.13	0.96	0.91	0.98	1.19	0.91	1.09	0.97	1.10	1.17	1.07	1.25	0.84	1.06	0.92
Hindlimb grip strength (kg)	Trial 1	0.47	0.43	0.53	0.62	0.42	0.53	0.51	0.54	0.69	0.32	0.55	0.50	0.45	0.51	0.55	0.44	0.45	0.44	0.59	0.53
	Trial 2	0.45	0.45	0.50	0.54	0.40	0.54	0.46	0.43	0.44	0.49	0.43	0.46	0.52	0.42	0.47	0.49	0.49	0.37	0.44	0.57
	Trial 3	0.55	0.48	0.40	0.48	0.49	0.53	0.44	0.45	0.60	0.49	0.44	0.58	0.44	0.37	0.49	0.46	0.41	0.43	0.47	0.49
	Mean	0.49	0.45	0.48	0.55	0.44	0.53	0.47	0.47	0.58	0.43	0.47	0.51	0.47	0.43	0.50	0.46	0.45	0.41	0.50	0.53
Additional observations		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Forelimb grip strength (kg)	Mean	1.07	1.04
	SD	0.12	0.13
	N	10	10
	P		
Hindlimb grip strength (kg)	Mean	0.49	0.47
	SD	0.05	0.04
	N	10	10
	P		

APPENDIX 4

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
1M	11	High	139	68	2	12	0	0	0	0	0	33	254
	12		114	60	18	10	0	0	0	0	0	1	203
	13		119	96	60	0	0	0	0	0	0	27	302
	14		110	65	18	0	0	0	0	0	0	0	193
	15		122	95	65	29	0	0	0	4	0	0	315
	16		92	82	42	4	16	18	5	0	0	22	281
	17		187	135	67	0	0	25	88	2	0	0	504
	18		124	22	15	0	0	0	0	0	0	0	161
	19		134	99	45	89	46	3	25	0	0	0	441
	20		141	134	93	36	44	17	45	3	0	0	513
	Mean		128.2	85.6	42.5	18.0	10.6	6.3	16.3	0.9	0.0	8.3	316.7
	SD		25.3	34.3	29.0	28.0	18.8	9.7	29.4	1.5	0.0	13.4	127.8
	N		10	10	10	10	10	10	10	10	10	10	10
	Method†		2.7	2.7	2.7	2.7	2.7	2.7	2.7	8	8	2.7	2.7
	P		0.851	0.891	0.727	0.357	0.812	0.191	0.567	0.138	0.261	0.853	0.683

† See Statistical Analysis Section in Experimental Procedure for explanation

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
1M	11	Low	204	183	101	36	4	0	0	0	0	43	571
	12		235	141	85	18	0	0	0	0	43	14	536
	13		243	195	114	51	1	0	0	0	0	23	627
	14		170	138	80	0	0	0	0	0	7	0	395
	15		240	249	153	100	8	0	0	27	3	3	783
	16		186	139	99	64	73	94	53	4	3	85	800
	17		180	209	99	25	2	42	104	22	2	50	735
	18		255	83	22	0	0	0	0	0	0	0	360
	19		235	116	69	56	35	32	47	14	0	1	605
	20		248	214	124	117	104	43	71	38	11	0	970
	Mean		219.6	166.7	94.6	46.7	22.7	21.1	27.5	10.5	6.9	21.9	638.2
	SD		31.4	51.3	34.9	39.4	37.0	31.7	38.5	14.0	13.2	28.9	188.4
	N		10	10	10	10	10	10	10	10	10	10	10
	Method†		2.7	2.7	2.7	2.7	2.7	2.7	2.7	8	8	2.7	2.7
	P		0.179	0.832	0.413	0.101	0.305	0.340	0.377	0.015	0.094	0.747	0.094

† See Statistical Analysis Section in Experimental Procedure for explanation

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
2M	31	High	123	68	34	37	35	0	0	0	0	0	297
	32		119	72	64	41	0	0	0	0	7	0	303
	33		132	86	21	37	8	12	0	0	0	26	322
	34		112	78	57	0	0	0	0	0	29	61	337
	35		131	45	0	0	0	0	0	0	0	0	176
	36		135	97	52	68	61	30	28	0	0	5	476
	37		103	81	50	13	0	0	0	0	0	0	247
	38		110	45	34	32	20	0	0	0	0	0	241
	39		153	78	69	66	58	27	30	2	0	0	483
	40		113	108	29	37	0	0	0	0	0	0	287
	Mean		123.1	75.8	41.0	33.1	18.2	6.9	5.8	0.2	3.6	9.2	316.9
	SD		14.9	20.0	21.3	23.6	24.6	12.0	12.2	0.6	9.2	19.9	97.4
	N		10	10	10	10	10	10	10	10	10	10	10
	P												

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
2M	31	Low	207	151	109	103	56	74	41	30	0	9	780
	32		228	178	146	64	72	33	3	23	45	0	792
	33		288	237	109	72	51	29	7	1	16	112	922
	34		194	209	93	10	1	0	0	0	95	166	768
	35		215	93	2	0	0	0	0	0	0	0	310
	36		255	142	169	100	105	76	121	24	0	11	1003
	37		316	206	105	39	0	0	3	0	0	0	669
	38		264	154	60	42	43	2	1	1	2	0	569
	39		304	203	123	122	82	75	70	19	5	0	1003
	40		281	185	43	109	7	0	22	0	0	0	647
	Mean		255.5	175.8	95.9	66.1	41.7	28.9	26.8	9.8	16.3	29.8	746.3
	SD		42.6	41.7	49.4	42.6	38.3	34.1	40.3	12.5	31.1	59.1	211.9
	N		10	10	10	10	10	10	10	10	10	10	10
	P									n.s.			

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
3M	1	High	100	58	61	17	0	0	0	0	0	42	278
	2		157	159	66	77	14	11	66	66	55	33	704
	3		92	97	19	0	0	0	0	0	0	16	224
	4		105	93	26	0	0	0	0	0	0	31	255
	5		109	110	69	29	25	8	6	7	24	0	387
	6		115	51	16	0	1	0	0	0	0	0	183
	7		152	59	8	20	51	20	2	34	0	0	346
	8		164	73	37	15	44	43	26	54	24	13	493
	9		95	38	27	0	0	0	2	0	0	0	162
	Mean		121.0	82.0	36.6	17.6	15.0	9.1	11.3	17.9	11.4	15.0	336.9
	SD		28.5	37.3	23.1	24.8	20.4	14.5	22.1	26.5	19.4	16.6	172.8
	N		9	9	9	9	9	9	9	9	9	9	9
	P												

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
3M	1	Low	167	102	89	62	45	10	7	12	5	42	541
	2		281	221	140	102	42	26	80	93	107	103	1195
	3		257	169	71	12	0	0	2	0	0	40	551
	4		236	174	80	20	10	1	13	3	0	117	654
	5		287	215	186	69	124	50	19	5	68	0	1023
	6		322	193	83	11	0	6	0	2	1	9	627
	7		256	127	29	23	48	27	44	50	0	0	604
	8		223	126	69	18	92	98	80	93	74	50	923
	9		241	97	126	12	0	3	12	17	57	0	565
	Mean		252.2	158.2	97.0	36.6	40.1	24.6	28.6	30.6	34.7	40.1	742.6
	SD		43.9	47.0	46.4	32.9	44.1	32.1	31.9	38.5	41.8	44.4	241.1
	N		9	9	9	9	9	9	9	9	9	9	9
	P									n.s.			

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
4M	21	High	121	126	78	2	11	0	0	0	12	45	395
	22		119	104	59	35	43	7	0	58	3	0	428
	23		169	76	3	0	0	0	0	0	0	0	248
	24		124	49	20	13	2	8	0	0	0	0	216
	25		98	65	46	9	34	13	2	11	0	0	278
	26		148	124	45	54	24	50	36	1	0	0	482
	27		96	62	44	59	7	28	12	0	0	4	312
	28		124	112	94	98	59	63	69	33	29	52	733
	29		115	60	44	16	1	29	32	14	0	0	311
	30		183	68	62	59	8	0	45	0	0	0	425
	Mean		129.7	84.6	49.5	34.5	18.9	19.8	19.6	11.7	4.4	10.1	382.8
	SD		28.5	28.9	26.2	32.2	20.4	22.3	24.5	19.4	9.4	20.3	150.4
	N		10	10	10	10	10	10	10	10	10	10	10
	P												

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
4M	21	Low	220	210	168	63	43	2	0	1	60	145	912
	22		239	144	92	89	83	14	16	73	54	0	804
	23		301	218	55	6	0	0	15	13	10	4	622
	24		169	111	99	56	15	49	45	53	70	13	680
	25		191	133	154	71	71	59	93	34	27	102	935
	26		279	115	122	96	71	155	74	27	20	13	972
	27		167	127	151	109	74	74	36	17	4	84	843
	28		249	131	129	95	127	47	77	125	112	63	1155
	29		262	188	121	71	42	65	63	111	8	6	937
	30		226	222	127	97	38	24	95	31	0	0	860
	Mean		230.3	159.9	121.8	75.3	56.4	48.9	51.4	48.5	36.5	43.0	872.0
	SD		45.0	44.5	33.3	29.8	36.6	45.6	34.0	42.0	36.4	52.0	150.9
	N		10	10	10	10	10	10	10	10	10	10	10
	P									<0.05			

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
1F	51	High	187	114	90	72	76	98	56	106	8	7	814
	52		210	147	48	36	53	47	92	55	70	33	791
	53		195	162	82	45	92	37	113	81	51	87	945
	54		162	140	112	64	61	65	60	58	40	18	780
	55		164	63	6	0	42	26	0	0	0	0	301
	56		282	154	156	86	112	94	72	80	33	91	1160
	57		126	151	48	42	50	20	15	19	16	1	488
	58		189	88	81	75	71	41	25	53	44	58	725
	59		187	90	49	68	0	0	0	21	55	48	518
	60		190	215	115	120	54	36	0	0	0	55	785
	Mean		189.2	132.4	78.7	60.8	61.1	46.4	43.3	47.3	31.7	39.8	730.7
	SD		40.1	44.4	42.9	32.4	30.2	31.2	41.2	36.2	24.5	33.6	243.8
	N		10	10	10	10	10	10	10	10	10	10	10
	Method†		2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7
	P		0.389	0.333	0.361	0.775	0.306	0.188	0.469	0.388	0.248	0.850	0.283

† See Statistical Analysis Section in Experimental Procedure for explanation

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
1F	51	Low	253	165	128	77	84	101	78	56	75	50	1067
	52		321	216	91	117	88	88	106	55	111	115	1308
	53		237	149	122	86	137	57	119	73	103	75	1158
	54		325	192	116	75	109	75	116	99	57	22	1186
	55		340	252	40	12	106	39	0	0	0	0	789
	56		243	142	140	192	171	107	207	81	50	47	1380
	57		214	110	121	93	90	67	114	92	57	62	1020
	58		217	168	137	133	140	119	100	125	122	99	1360
	59		189	126	95	165	9	1	3	46	228	130	992
	60		207	146	127	74	98	150	6	0	11	36	855
	Mean		254.6	166.6	111.7	102.4	103.2	80.4	84.9	62.7	81.4	63.6	1111.5
	SD		54.5	42.9	29.8	51.5	43.2	42.6	65.6	40.5	65.2	41.4	204.0
	N		10	10	10	10	10	10	10	10	10	10	10
	Method†		3.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	8	2.7	2.7
	P		0.880	0.452	0.889	0.765	0.354	0.442	0.249	0.312	0.880	0.678	0.394

† See Statistical Analysis Section in Experimental Procedure for explanation

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
2F	41	High	134	54	50	39	17	14	24	20	0	0	352
	42		264	206	110	176	35	128	48	65	24	108	1164
	43		175	107	71	39	37	1	61	15	14	36	556
	44		127	89	58	37	30	33	15	23	5	0	417
	45		232	180	110	83	64	28	37	100	19	60	913
	46		147	37	35	30	21	0	0	0	27	55	352
	47		179	154	89	60	45	7	14	25	23	23	619
	48		155	69	13	35	0	21	1	39	32	49	414
	49		206	102	89	44	73	39	31	8	23	45	660
	50		201	129	107	59	81	57	64	20	8	64	790
	Mean		182.0	112.7	73.2	60.2	40.3	32.8	29.5	31.5	17.5	44.0	623.7
	SD		44.1	54.8	33.7	43.7	25.8	37.9	23.0	29.9	10.4	32.1	267.4
	N		10	10	10	10	10	10	10	10	10	10	10
	P												

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
2F	41	Low	304	125	143	249	112	88	100	92	12	15	1240
	42		299	206	128	53	68	91	44	116	155	70	1230
	43		227	99	91	89	114	33	129	71	97	99	1049
	44		304	163	72	68	54	81	48	96	13	2	901
	45		384	325	225	97	108	127	70	122	53	109	1620
	46		158	69	53	84	50	7	7	78	228	95	829
	47		175	109	127	109	105	55	32	41	40	68	861
	48		274	184	108	75	12	95	19	106	95	62	1030
	49		232	161	110	131	106	124	96	10	70	81	1121
	50		255	206	151	168	82	131	128	112	51	118	1402
	Mean		261.2	164.7	120.8	112.3	81.1	83.2	67.3	84.4	81.4	71.9	1128.3
	SD		67.1	72.8	47.7	58.3	34.3	41.1	44.1	35.7	67.0	38.1	251.3
	N		10	10	10	10	10	10	10	10	10	10	10
	P												

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
3F	61	High	119	73	24	22	10	19	19	0	0	0	286
	62		192	96	32	29	2	0	0	0	44	4	399
	63		176	57	43	41	14	0	4	19	0	2	356
	64		162	91	44	51	33	0	0	0	23	2	406
	65		91	45	34	57	21	0	19	29	33	0	329
	66		163	110	61	38	22	0	0	0	13	53	460
	67		217	181	125	126	91	40	97	115	22	87	1101
	68		155	65	64	37	46	66	31	13	25	53	555
	69		178	144	97	58	81	44	65	27	24	91	809
	70		151	44	33	42	40	0	2	34	25	22	393
	Mean		160.4	90.6	55.7	50.1	36.0	16.9	23.7	23.7	20.9	31.4	509.4
	SD		35.6	44.4	32.4	29.0	29.6	24.5	32.9	34.7	13.6	36.7	255.3
	N		10	10	10	10	10	10	10	10	10	10	10
	P												

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
3F	61	Low	263	117	90	110	63	134	42	35	40	62	956
	62		267	145	99	34	7	5	8	12	67	50	694
	63		208	121	90	78	38	4	20	53	0	39	651
	64		209	106	81	38	33	13	0	0	52	39	571
	65		247	104	71	71	28	2	40	78	70	0	711
	66		254	172	122	135	97	12	0	0	409	115	1316
	67		241	137	152	154	151	91	55	108	185	45	1319
	68		236	124	129	98	85	72	120	31	26	56	977
	69		277	194	184	142	121	152	76	104	80	74	1404
	70		220	103	87	69	74	58	46	74	58	36	825
	Mean		242.2	132.3	110.5	92.9	69.7	54.3	40.7	49.5	98.7	51.6	942.4
	SD		24.1	30.5	35.8	42.2	45.1	56.5	37.4	40.4	119.4	29.7	306.9
	N		10	10	10	10	10	10	10	10	10	10	10
	P												

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
4F	71	High	202	85	41	65	41	14	20	0	31	61	560
	72		155	92	78	44	29	97	3	6	55	63	622
	73		217	143	62	65	49	34	29	10	35	24	668
	74		154	71	93	78	57	50	19	25	36	23	606
	75		226	91	74	45	9	34	27	49	19	5	579
	76		146	63	28	60	18	3	0	0	0	54	372
	77		232	248	179	140	118	56	50	58	43	54	1178
	78		165	158	175	129	107	116	46	59	33	29	1017
	79		156	107	74	42	67	14	56	65	33	71	685
	80		161	91	73	5	3	26	0	1	0	1	361
	Mean		181.4	114.9	87.7	67.3	49.8	44.4	25.0	27.3	28.5	38.5	664.8
	SD		33.8	55.4	50.6	40.6	38.9	36.8	20.7	27.4	17.6	25.2	255.9
	N		10	10	10	10	10	10	10	10	10	10	10
	P												

APPENDIX 4 - continued

Motor activity - individual scores (beam breaks) during Week 12 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal number	Beam level	Time (minutes) 6	12	18	24	30	36	42	48	54	60	Total
4F	71	Low	216	104	33	98	61	28	43	11	35	91	720
	72		285	219	107	68	102	89	30	55	131	81	1167
	73		332	154	90	72	92	105	87	167	93	41	1233
	74		187	117	53	67	70	73	73	33	40	59	772
	75		369	242	72	63	68	43	133	89	45	19	1143
	76		197	92	100	72	77	52	10	1	19	145	765
	77		269	166	222	175	140	68	64	91	119	46	1360
	78		289	88	74	35	59	38	23	32	88	18	744
	79		293	208	206	221	157	110	113	89	44	113	1554
	80		223	117	92	12	6	86	39	10	6	49	640
	Mean		266.0	150.7	104.9	88.3	83.2	69.2	61.5	57.8	62.0	66.2	1009.8
	SD		59.6	56.1	61.7	63.2	42.9	28.5	40.2	51.6	42.8	41.1	319.5
	N		10	10	10	10	10	10	10	10	10	10	10
	P												

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
1M	11	265	304	344	372	398	412	429	442	462	471
	12	293	353	391	426	465	480	508	539	576	599
	13	261	316	360	389	419	437	461	489	507	519
	14	243	293	326	352	373	392	405	421	436	444
	15	255	312	360	391	423	449	467	490	509	520
	16	262	325	376	411	445	468	486	504	515	534
	17	245	275	304	317	341	359	374	384	394	407
	18	242	281	314	332	351	368	378	380	387	390
	19	257	301	346	371	396	414	437	448	460	471
	20	253	293	336	356	394	409	426	441	455	472
	Mean	257	305	346	372	400	419	437	454	470	483
	SD	14.8	22.7	27.3	33.9	39.0	40.2	44.2	51.7	58.1	62.3
	N	10	10	10	10	10	10	10	10	10	10
	Method †	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7
	P	0.634	0.403	0.589	0.678	0.635	0.652	0.644	0.811	0.808	0.817

† See Statistical Analysis Section in Experimental Procedure for explanation

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 10	Week 11	Week 12	Week 13
1M	11	483	489	500	500
	12	614	630	642	639
	13	528	543	554	560
	14	450	463	472	475
	15	536	551	560	562
	16	547	558	560	556
	17	415	425	435	436
	18	395	405	410	418
	19	479	493	494	496
	20	481	497	502	509
	Mean	493	505	513	515
	SD	65.5	67.2	68.1	66.1
	N	10	10	10	10
	Method †	2.7	2.7	2.7	2.7
	P	0.791	0.847	0.811	0.804

† See Statistical Analysis Section in Experimental Procedure for explanation

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
2M	31	249	295	345	378	413	436	464	483	506	523
	32	279	339	404	458	501	537	570	580	601	614
	33	254	299	337	360	384	401	423	436	451	458
	34	256	310	356	388	422	438	462	476	491	503
	35	242	286	336	363	394	416	436	449	476	489
	36	251	289	337	376	409	430	448	452	464	479
	37	270	316	373	410	445	461	484	496	505	519
	38	238	275	319	353	378	402	430	457	477	496
	39	237	281	326	355	377	404	420	441	458	473
	40	264	299	331	360	379	392	406	414	428	439
	Mean	254	299	346	380	410	432	454	468	486	499
	SD	13.7	18.8	25.4	32.6	39.1	42.8	47.0	46.1	47.2	48.0
	N	10	10	10	10	10	10	10	10	10	10
	P										

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 10	Week 11	Week 12	Week 13
2M	31	549	559	569	566
	32	635	638	664	651
	33	464	465	476	477
	34	509	511	517	521
	35	500	509	516	525
	36	487	495	501	506
	37	523	534	549	555
	38	508	520	529	535
	39	481	489	497	492
	40	439	452	456	457
	Mean	510	517	527	528
	SD	53.6	52.7	58.3	54.8
	N	10	10	10	10
	P				

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
3M	1	253	302	359	402	445	479	508	533	546	564
	2	238	273	302	325	345	369	388	396	409	419
	3	259	303	350	383	416	432	449	458	470	482
	4	259	288	332	355	375	397	408	421	436	449
	5	247	285	321	352	375	398	416	431	450	456
	6	248	294	342	377	405	430	448	455	469	478
	7	240	289	350	388	437	461	485	506	520	541
	8	230	266	299	325	353	378	404	427	446	456
	9	259	307	366	405	447	471	500	520	539	552
	10	272	325	381							
	Mean	250	293	340	368	400	424	445	461	476	489
	SD	12.4	17.2	26.6	30.4	39.3	40.9	44.4	48.0	48.1	51.5
	N	10	10	10	9	9	9	9	9	9	9
	P										

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 10	Week 11	Week 12	Week 13
3M	1	587	607	615	615
	2	424	437	441	438
	3	490	506	508	502
	4	450	463	464	464
	5	466	474	486	485
	6	485	497	505	504
	7	552	570	583	579
	8	464	479	493	492
	9	563	573	582	573
	10				
	Mean	498	512	520	517
	SD	56.2	58.0	59.8	58.9
	N	9	9	9	9
	P				

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
4M	21	232	268	297	326	350	361	379	405	418	425
	22	260	307	355	390	416	438	459	480	496	501
	23	244	273	299	317	334	346	358	362	376	383
	24	261	308	353	386	416	436	460	474	483	499
	25	266	299	329	359	381	417	439	469	494	509
	26	251	292	332	352	373	389	407	428	444	460
	27	252	299	342	378	402	426	452	469	487	507
	28	257	305	363	408	444	474	505	534	554	571
	29	245	283	318	346	371	387	401	418	428	442
	30	248	293	336	367	394	416	436	445	460	473
	Mean	252	293	332	363	388	409	430	448	464	477
	SD	10.1	14.0	22.6	28.7	33.1	38.6	43.5	47.8	50.0	52.5
	N	10	10	10	10	10	10	10	10	10	10
	P										

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 10	Week 11	Week 12	Week 13
4M	21	429	434	441	443
	22	513	526	532	529
	23	383	391	399	409
	24	512	521	528	529
	25	508	520	528	531
	26	460	469	474	477
	27	517	534	545	547
	28	586	604	611	609
	29	448	457	464	461
	30	481	489	490	483
	Mean	484	494	501	502
	SD	56.7	60.0	60.1	58.4
	N	10	10	10	10
	P				

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
1F	51	177	210	235	246	247	272	279	285	281	293
	52	171	192	217	236	251	261	269	282	282	291
	53	182	196	219	244	266	274	293	307	324	315
	54	181	209	234	245	274	282	291	290	304	313
	55	150	179	198	213	222	236	248	255	258	262
	56	172	200	220	240	249	260	276	283	284	291
	57	172	197	230	250	264	271	286	294	302	300
	58	158	184	202	210	236	244	252	243	256	261
	59	172	207	226	253	276	281	288	302	309	312
	60	180	202	220	229	246	257	265	263	278	287
	Mean	172	197	220	236	253	264	275	280	288	292
	SD	10.2	10.2	12.2	15.1	17.2	15.3	16.0	20.6	22.1	19.3
	N	10	10	10	10	10	10	10	10	10	10
	Method †	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7
	P	0.586	0.692	0.456	0.579	0.541	0.428	0.315	0.476	0.453	0.476

† See Statistical Analysis Section in Experimental Procedure for explanation

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 10	Week 11	Week 12	Week 13
1F	51	301	306	300	311
	52	292	296	293	294
	53	326	336	334	326
	54	314	318	325	331
	55	268	272	269	278
	56	299	300	299	302
	57	311	316	320	320
	58	266	261	268	275
	59	310	320	325	323
	60	286	288	291	297
	Mean	297	301	302	306
	SD	19.6	22.9	23.2	19.8
	N	10	10	10	10
	Method †	2.7	2.7	2.7	2.7
	P	0.441	0.423	0.473	0.546

† See Statistical Analysis Section in Experimental Procedure for explanation

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
2F	41	169	189	208	229	238	249	260	270	271	276
	42	182	206	222	229	247	263	271	275	282	289
	43	155	174	188	211	223	230	235	245	251	259
	44	151	177	199	218	220	239	249	257	261	271
	45	163	188	209	224	233	251	261	269	270	279
	46	164	186	205	218	233	251	260	264	277	289
	47	180	201	228	254	270	274	297	304	314	306
	48	171	190	214	234	245	245	264	273	280	283
	49	182	207	230	244	265	276	291	297	308	316
	50	182	205	224	249	264	271	271	282	295	296
	Mean	170	192	213	231	244	255	266	274	281	286
	SD	11.7	11.9	13.4	14.4	17.8	15.7	18.3	17.5	19.7	16.7
	N	10	10	10	10	10	10	10	10	10	10
	P										

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 10	Week 11	Week 12	Week 13
2F	41	280	285	296	300
	42	295	298	299	291
	43	261	269	271	272
	44	272	277	273	284
	45	285	294	283	292
	46	292	293	294	302
	47	317	329	330	347
	48	296	303	292	303
	49	322	327	333	341
	50	291	302	312	310
	Mean	291	298	298	304
	SD	18.7	19.2	21.3	23.5
	N	10	10	10	10
	P				

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
3F	61	157	177	189	206	227	229	230	237	247	255
	62	157	182	202	210	218	232	247	250	250	254
	63	183	209	233	244	270	284	294	294	314	325
	64	165	191	214	242	267	261	268	284	295	291
	65	169	187	206	220	231	241	258	261	264	269
	66	177	208	232	253	274	283	288	293	297	306
	67	158	188	207	231	245	259	263	276	284	295
	68	168	191	206	216	227	238	254	262	261	270
	69	169	197	224	241	251	263	278	289	283	296
	70	181	210	227	240	264	277	283	287	293	303
	Mean	168	194	214	230	247	257	266	273	279	286
	SD	9.5	11.6	14.7	16.1	20.9	20.8	20.1	19.8	22.3	23.4
	N	10	10	10	10	10	10	10	10	10	10
	P										

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 10	Week 11	Week 12	Week 13
3F	61	244	257	264	264
	62	261	268	260	266
	63	328	329	341	351
	64	302	308	315	305
	65	278	281	278	284
	66	307	310	313	309
	67	294	304	313	311
	68	276	281	273	274
	69	307	313	306	314
	70	305	309	314	317
	Mean	290	296	298	300
	SD	25.0	23.0	27.1	27.3
	N	10	10	10	10
	P				

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
4F	71	180	206	232	246	260	274	292	300	298	313
	72	186	211	242	262	274	297	319	332	330	350
	73	170	188	214	228	250	261	276	282	289	302
	74	152	168	196	204	218	231	242	239	253	250
	75	191	215	246	266	281	282	299	312	320	316
	76	169	194	210	233	250	259	272	281	288	292
	77	167	187	199	222	230	241	249	253	259	261
	78	168	192	209	220	242	251	258	260	272	278
	79	187	214	245	269	287	298	318	332	338	338
	80	178	203	225	243	261	276	286	285	296	308
	Mean	175	198	222	239	255	267	281	287	294	301
	SD	11.9	15.0	18.9	21.6	21.9	22.3	27.0	31.7	28.6	31.7
	N	10	10	10	10	10	10	10	10	10	10
	P										

Bodyweights - individual values (g)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Week 10	Week 11	Week 12	Week 13
4F	71	309	328	330	328
	72	359	369	367	367
	73	303	300	310	315
	74	265	269	263	272
	75	326	346	345	341
	76	302	307	314	309
	77	271	275	268	280
	78	272	281	284	272
	79	349	362	362	358
	80	308	302	307	323
	Mean	306	314	315	316
	SD	31.9	36.0	36.4	33.9
	N	10	10	10	10
	P				

Food consumption - individual values (g/animal/week)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Cage Number	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11
1M	3	204	197	198	192	197	256	201	205	202	201	202
	4	185	182	179	180	184	184	184	186	184	181	184
2M	7	192	198	199	195	200	202	204	201	198	198	195
	8	187	196	197	188	193	192	191	192	190	189	189
3M	1	185	183	181	182	194	193	189	186	191	184	184
	2	195	195	184	200	200	200	203	202	195	189	189
4M	5	176	170	172	170	179	180	178	182	180	174	171
	6	183	186	186	179	191	193	189	192	191	188	182

Food consumption - individual values (g/animal/week)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Cage Number	Week 12	Week 13
1M	3	176	191
	4	176	174
2M	7	196	178
	8	184	178
3M	1	178	174
	2	186	181
4M	5	167	163
	6	177	170

Food consumption - individual values (g/animal/week)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Cage Number	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11
1F	11	127	134	137	142	147	150	144	138	139	135	136
	12	126	132	136	136	137	142	139	131	132	129	130
2F	9	119	124	132	132	138	139	141	134	134	127	137
	10	127	132	143	142	142	151	144	141	138	142	142
3F	13	118	122	127	131	126	132	129	126	126	122	127
	14	122	124	130	130	133	128	133	125	129	126	127
4F	15	128	134	137	141	148	150	149	144	140	144	147
	16	126	129	142	139	146	145	144	137	135	134	138

Food consumption - individual values (g/animal/week)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Cage Number	Week 12	Week 13
1F	11	132	128
	12	126	122
2F	9	127	122
	10	139	129
3F	13	119	117
	14	124	122
4F	15	133	130
	16	130	119

Water consumption - individual values (ml/animal/day)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Number	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11
1M	3	34	35	34	33	36	36	38	39	40	43	41
	4	37	34	37	35	39	35	36	38	43	41	41
2M	7	34	33	36	34	35	34	32	35	34	36	42
	8	31	31	31	30	34	31	31	30	32	35	38
3M	1	37	35	40	36	40	42	40	43	44	41	43
	2	32	33	32	31	32	32	32	32	33	31	31
4M	5	34	34	32	29	33	31	32	42	36	32	33
	6	34	32	36	35	40	33	39	38	38	38	37

Water consumption - individual values (ml/animal/day)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Number	Week 12	Week 13
1M	3	38	37
	4	39	38
2M	7	35	31
	8	34	31
3M	1	40	39
	2	33	27
4M	5	30	30
	6	32	36

Water consumption - individual values (ml/animal/day)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Number	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11
1F	11	25	27	28	26	29	28	30	27	32	30	32
	12	23	23	23	23	26	25	27	23	30	27	25
2F	9	23	28	28	26	30	29	28	30	31	30	29
	10	23	23	25	25	25	23	28	25	26	25	24
3F	13	20	24	22	23	26	24	26	27	25	24	23
	14	21	22	23	22	24	21	23	23	26	23	22
4F	15	25	25	28	25	32	28	32	30	29	27	29
	16	26	24	29	27	34	26	29	40	31	31	29

Water consumption - individual values (ml/animal/day)

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Number	Week 12	Week 13
1F	11	26	25
	12	22	26
2F	9	29	28
	10	23	23
3F	13	24	24
	14	21	21
4F	15	24	27
	16	27	27

APPENDIX 8

Ophthalmoscopy - individual observations before commencement of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

		Group and sex/Animal number							
Structure	Observation	1M	2M	3M	4M	1F	2F	3F	4F
Conjunctiva	(Slightly) prominent aqueous vein	16R							
Cornea/sclera	Slight superficial corneal opacity			2 5L		60			
	Short linear superficial corneal opacity		38	2R					
	Corneal opacity spot	12L							
Lens	Faint nuclear opacity	13L	39						
	Nuclear opacity spot								73L
	Posterior polar opacity		38L						
	Focal posterior capsular opacity						48R		
Fundus	(Radial) folds		39L						78L
	Rosette cluster		39L						
Number of animals examined :		10	10	10	10	10	10	10	10

L Left R Right

APPENDIX 8 - continued

Ophthalmoscopy - individual observations during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

		Group and sex/Animal number			
Structure	Observation	1M	4M	1F	4F
Conjunctiva	(Slightly) prominent aqueous vein	16R			
Cornea/sclera	Slight superficial corneal opacity			53L 60R	80R
	Short linear superficial corneal opacity				
	Corneal opacity spot	12L			
Iris	Focal haemorrhage	20L			
Lens	(Faint) nuclear opacity	11R 13L	23R 28R	52R	
	(Faint) nuclear opacity spot	15L	24R 30L	51	73L
	Not observable	20L			
Fundus	Hyperreflection bands			58L	
	Short radial folds				78L
	Not observable	20L			
Other	Hyphaema and aqueous flare	20L			
Number of animals examined :		10	10	10	10

L Left R Right

APPENDIX 9

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Hct L/L	Hb g/dL	RBC $x10^{12}$/L	MCH pg	MCHC g/dL	MCV fL
1M	11	0.459	15.6	8.33	18.8	34.0	55.1
	12	0.450	15.5	7.96	19.5	34.5	56.5
	13	0.462	15.5	8.54	18.2	33.6	54.1
	14	0.493	16.4	9.11	18.0	33.4	54.1
	15	0.472	16.1	8.49	19.0	34.2	55.6
	16	0.485	16.1	8.86	18.2	33.3	54.8
	17	0.468	15.5	8.80	17.6	33.1	53.1
	18	0.480	16.4	8.59	19.1	34.1	56.0
	19	0.471	15.8	8.41	18.8	33.6	56.1
	20	0.476	15.8	8.53	18.5	33.2	55.8
	Mean	0.472	15.9	8.56	18.6	33.7	55.1
	SD	0.0127	0.36	0.315	0.58	0.47	1.09
	n	10	10	10	10	10	10
	Method †	2.7	8	2.7	2.7	2.7	2.7
	P	0.035	0.121	0.282	0.624	0.762	0.357

† See statistical analysis section in Experimental Procedure for explanation

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	WBC $x10^9$/L	N $x10^9$/L	L $x10^9$/L	E $x10^9$/L	B $x10^9$/L	M $x10^9$/L	LUC $x10^9$/L
1M	11	9.18	1.13	7.64	0.12	0.02	0.24	0.03
	12	11.79	1.19	10.13	0.15	0.04	0.23	0.04
	13	15.03	1.82	12.59	0.16	0.10	0.32	0.04
	14	8.93	1.67	6.73	0.10	0.05	0.32	0.06
	15	13.05	1.07	11.52	0.13	0.07	0.22	0.05
	16	14.86	2.31	11.84	0.14	0.06	0.46	0.04
	17	12.22	1.03	10.90	0.08	0.04	0.12	0.05
	18	12.69	2.24	10.00	0.08	0.07	0.28	0.02
	19	12.96	1.48	11.05	0.16	0.08	0.19	0.02
	20	15.35	0.87	14.02	0.16	0.08	0.18	0.04
	Mean	12.61	1.48	10.64	0.13	0.06	0.26	0.04
	SD	2.233	0.513	2.177	0.032	0.024	0.095	0.013
	n	10	10	10	10	10	10	10
	Method †	2.7	3.7	2.7	2.7	2.7	2.7	3.7
	P	0.129	0.745	0.089	0.426	0.316	0.777	0.782

† See statistical analysis section in Experimental Procedure for explanation

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Plt x10^9/L	PT sec	APTT sec	Aniso-cytosis	Micro-cytosis	Macro-cytosis	Hypo-chromasia	Hyper-chromasia
1M	11	1017	14.6	17.4	-	-	-	+	-
	12	1511	16.0	14.1	-	-	-	++	-
	13	1212	14.5	14.8	-	+	-	++	-
	14	1068	15.1	18.2	-	-	-	+	-
	15	1010	14.2	17.8	-	-	-	+	-
	16	1126	16.1	15.5	-	-	-	++	-
	17	1064	15.0	17.7	-	-	-	++	-
	18	908	16.8	13.9	-	-	-	-	-
	19	1152	15.4	16.0	-	-	-	++	-
	20	1283	15.7	17.7	-	-	-	++	-
	Mean	1135	15.3	16.3					
	SD	170.0	0.82	1.65					
	n	10	10	10					
	Method †	8	2.7	2.7					
	P	0.980	0.436	0.162					

† See statistical analysis section in Experimental Procedure for explanation

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Hct L/L	Hb g/dL	RBC $x10^{12}$/L	MCH pg	MCHC g/dL	MCV fL
2M	31	0.474	15.9	8.60	18.5	33.5	55.1
	32	0.503	16.8	9.24	18.1	33.3	54.4
	33	0.434	14.8	7.86	18.8	34.2	55.2
	34	0.461	15.4	8.33	18.5	33.4	55.4
	35	0.482	16.2	8.75	18.5	33.5	55.0
	36	0.467	15.6	8.90	17.6	33.5	52.4
	37	0.447	14.5	8.23	17.7	32.5	54.3
	38	0.476	16.2	8.48	19.1	34.0	56.1
	39	0.478	16.5	8.96	18.5	34.6	53.4
	40	0.447	15.0	8.07	18.6	33.6	55.4
	Mean	0.467	15.7	8.54	18.4	33.6	54.7
	SD	0.0203	0.76	0.431	0.47	0.57	1.09
	n	10	10	10	10	10	10
	P	0.522					

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	WBC $x10^9$/L	N $x10^9$/L	L $x10^9$/L	E $x10^9$/L	B $x10^9$/L	M $x10^9$/L	LUC $x10^9$/L
2M	31	16.60	3.36	12.31	0.29	0.09	0.43	0.11
	32	9.91	1.56	7.85	0.15	0.06	0.27	0.02
	33	10.08	1.39	8.34	0.09	0.03	0.22	0.02
	34	13.72	1.38	11.88	0.19	0.08	0.16	0.04
	35	11.69	1.38	9.77	0.14	0.08	0.29	0.02
	36	13.98	1.20	12.38	0.13	0.06	0.18	0.04
	37	13.20	1.51	11.08	0.16	0.08	0.32	0.04
	38	19.60	4.82	13.95	0.15	0.12	0.51	0.04
	39	12.14	1.25	10.38	0.11	0.08	0.30	0.02
	40	14.29	0.94	12.88	0.14	0.09	0.19	0.06
	Mean	13.52	1.88	11.08	0.16	0.08	0.29	0.04
	SD	2.935	1.227	1.985	0.055	0.024	0.112	0.028
	n	10	10	10	10	10	10	10
	P							

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Plt x10^9/L	PT sec	APTT sec	Aniso-cytosis	Micro-cytosis	Macro-cytosis	Hypo-chromasia	Hyper-chromasia
2M	31	1233	14.6	17.5	-	-	-	+	-
	32	1344	15.8	16.2	-	-	-	++	-
	33	1378	16.2	13.7	-	-	-	+++	-
	34	735	14.7	12.4	-	-	-	+	-
	35	1100	16.4	14.9	-	-	-	+	-
	36	1138	15.9	16.9	-	+	-	+	-
	37	1473	15.7	13.7	-	+	-	+++	-
	38	991	15.6	16.8	-	-	-	+	-
	39	868	16.6	16.7	-	-	-	+	-
	40	1124	16.2	15.1	-	-	-	+	-
	Mean	1138	15.8	15.4					
	SD	230.9	0.67	1.70					
	n	10	10	10					
	P								

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Hct L/L	Hb g/dL	RBC $x10^{12}$/L	MCH pg	MCHC g/dL	MCV fL
3M	1	0.462	15.4	8.21	18.8	33.4	56.2
	2	0.439	15.3	8.22	18.6	34.8	53.5
	3	0.466	15.7	8.85	17.7	33.7	52.7
	4	0.480	16.0	8.59	18.7	33.4	55.8
	5	0.455	15.2	8.45	18.0	33.5	53.9
	6	0.465	15.7	8.57	18.3	33.7	54.3
	7	0.456	15.6	8.54	18.3	34.3	53.4
	8	0.469	16.0	8.66	18.5	34.1	54.1
	9	0.468	15.8	8.46	18.7	33.7	55.4
	Mean	0.462	15.6	8.51	18.4	33.8	54.4
	SD	0.0114	0.29	0.203	0.36	0.47	1.19
	n	9	9	9	9	9	9
	P	0.217					

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	WBC $x10^9$/L	N $x10^9$/L	L $x10^9$/L	E $x10^9$/L	B $x10^9$/L	M $x10^9$/L	LUC $x10^9$/L
3M	1	13.85	3.19	10.00	0.19	0.06	0.38	0.03
	2	9.09	1.22	7.54	0.07	0.03	0.21	0.03
	3	10.20	1.84	7.83	0.08	0.07	0.36	0.03
	4	10.09	1.12	8.62	0.09	0.04	0.19	0.03
	5	11.86	1.66	9.67	0.08	0.07	0.33	0.04
	6	12.02	1.35	10.16	0.12	0.07	0.28	0.03
	7	9.99	2.02	7.48	0.10	0.08	0.29	0.03
	8	8.84	1.16	7.26	0.12	0.06	0.22	0.02
	9	13.12	1.36	11.12	0.21	0.06	0.32	0.06
	Mean	11.01	1.66	8.85	0.12	0.06	0.29	0.03
	SD	1.775	0.654	1.418	0.050	0.016	0.068	0.011
	n	9	9	9	9	9	9	9
	P							

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Plt x10^9/L	PT sec	APTT sec	Aniso-cytosis	Micro-cytosis	Macro-cytosis	Hypo-chromasia	Hyper-chromasia
3M	1	1206	14.9	20.9	-	-	-	+	-
	2	1010	16.2	16.9	-	-	-	-	-
	3	1196	15.9	15.6	-	+	-	+	-
	4	1211	16.3	14.1	-	-	-	-	-
	5	1229	16.6	15.7	-	+	-	++	-
	6	1080	15.7	13.3	-	-	-	-	-
	7	1019	15.1	15.0	-	-	-	+	-
	8	1072	14.3	17.9	-	-	-	-	-
	9	1080	17.3	18.4	-	-	-	+	-
	Mean	1123	15.8	16.4					
	SD	87.4	0.93	2.36					
	n	9	9	9					
	P								

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Hct L/L	Hb g/dL	RBC $x10^{12}$/L	MCH pg	MCHC g/dL	MCV fL
4M	21	0.440	14.8	8.12	18.2	33.6	54.2
	22	0.459	15.6	8.41	18.5	33.9	54.6
	23	0.448	15.2	8.17	18.6	33.8	54.9
	24	0.443	14.8	8.12	18.2	33.4	54.5
	25	0.479	16.1	9.14	17.6	33.6	52.4
	26	0.421	13.9	7.78	17.9	33.1	54.2
	27	0.432	14.6	8.22	17.7	33.8	52.5
	28	0.452	15.5	8.26	18.7	34.3	54.7
	29	0.478	16.3	8.55	19.1	34.2	55.9
	30	0.450	15.0	8.16	18.4	33.4	55.1
	Mean	0.450	15.2	8.29	18.3	33.7	54.3
	SD	0.0183	0.72	0.359	0.47	0.37	1.09
	n	10	10	10	10	10	10
	P	0.006					

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	WBC $x10^9$/L	N $x10^9$/L	L $x10^9$/L	E $x10^9$/L	B $x10^9$/L	M $x10^9$/L	LUC $x10^9$/L
4M	21	8.50	1.63	6.51	0.09	0.03	0.21	0.03
	22	13.44	1.30	11.59	0.18	0.06	0.28	0.04
	23	7.68	0.99	6.38	0.06	0.03	0.21	0.01
	24	14.69	3.10	10.90	0.19	0.06	0.40	0.04
	25	12.23	1.28	10.24	0.22	0.10	0.33	0.07
	26	13.74	0.69	12.65	0.06	0.09	0.20	0.04
	27	11.73	1.37	9.89	0.13	0.09	0.24	0.02
	28	14.06	1.27	12.21	0.23	0.05	0.25	0.04
	29	10.58	1.51	8.56	0.15	0.07	0.26	0.03
	30	10.11	1.82	7.91	0.10	0.06	0.21	0.02
	Mean	11.68	1.50	9.68	0.14	0.06	0.26	0.03
	SD	2.407	0.646	2.262	0.063	0.024	0.064	0.016
	n	10	10	10	10	10	10	10
	P							

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Plt x10^9/L	PT sec	APTT sec	Aniso-cytosis	Micro-cytosis	Macro-cytosis	Hypo-chromasia	Hyper-chromasia
4M	21	1065	14.8	20.4	-	-	-	++	-
	22	1146	16.0	15.6	-	-	-	+	-
	23	1199	15.5	21.6	-	-	-	+	-
	24	1212	16.5	15.8	-	+	-	++	-
	25	877	15.4	19.5	-	-	-	++	-
	26	1199	16.7	13.2	-	-	-	++	-
	27	1055	16.8	20.6	-	-	-	-	-
	28	1300	15.4	16.2	-	-	-	+	-
	29	1073	14.8	19.8	-	-	-	-	-
	30	1105	17.1	14.4	-	-	-	+	-
	Mean	1123	15.9	17.7					
	SD	116.7	0.84	2.98					
	n	10	10	10					
	P								

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Hct L/L	Hb g/dL	RBC $x10^{12}$/L	MCH pg	MCHC g/dL	MCV fL
1F	51	0.447	15.0	8.02	18.7	33.5	55.8
	52	0.441	15.3	8.06	19.0	34.7	54.8
	53	0.425	14.6	7.52	19.5	34.4	56.6
	54	0.450	15.7	7.93	19.8	35.0	56.7
	55	0.450	15.7	8.10	19.3	34.8	55.6
	56	0.423	14.3	7.80	18.4	33.8	54.3
	57	0.412	13.9	6.94	20.0	33.8	59.4
	58	0.456	15.9	8.11	19.6	34.9	56.2
	59	0.435	15.0	8.01	18.7	34.5	54.3
	60	0.451	15.3	7.96	19.2	33.9	56.7
	Mean	0.439	15.1	7.85	19.2	34.3	56.0
	SD	0.0147	0.65	0.363	0.52	0.54	1.50
	n	10	10	10	10	10	10
	Method †	2.7	2.7	2.7	2.7	2.7	2.7
	P	0.318	0.551	0.513	0.938	0.552	0.664

† See statistical analysis section in Experimental Procedure for explanation

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	WBC $x10^9$/L	N $x10^9$/L	L $x10^9$/L	E $x10^9$/L	B $x10^9$/L	M $x10^9$/L	LUC $x10^9$/L
1F	51	10.96	1.05	9.54	0.12	0.04	0.17	0.03
	52	9.49	0.91	8.31	0.12	0.02	0.12	0.02
	53	9.38	0.80	8.27	0.10	0.02	0.15	0.04
	54	8.13	0.58	7.26	0.07	0.02	0.16	0.02
	55	6.07	1.16	4.67	0.05	0.02	0.16	0.01
	56	7.70	0.49	6.88	0.13	0.03	0.16	0.02
	57	12.45	1.42	10.62	0.11	0.06	0.21	0.03
	58	4.61	0.32	4.07	0.11	0.01	0.08	0.01
	59	9.32	0.63	8.35	0.14	0.02	0.14	0.04
	60	6.54	0.38	5.94	0.08	0.02	0.11	0.02
	Mean	8.47	0.77	7.39	0.10	0.03	0.15	0.02
	SD	2.351	0.359	2.066	0.028	0.014	0.036	0.011
	n	10	10	10	10	10	10	10
	Method †	2.7	3.7	2.7	3.7	2.7	4.7	2.7
	P	0.373	0.073	0.419	0.157	0.944	0.127	0.820

† See statistical analysis section in Experimental Procedure for explanation

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Plt x10^9/L	PT sec	APTT sec	Aniso-cytosis	Micro-cytosis	Macro-cytosis	Hypo-chromasia	Hyper-chromasia
1F	51	1108	14.7	17.5	-	-	-	-	-
	52	934	15.9	16.0	-	-	-	-	-
	53	1103	14.8	12.0	-	-	-	-	-
	54	942	15.8	12.0	-	-	-	-	-
	55	888	15.3	16.3	-	-	-	-	-
	56	1226	14.5	13.5	-	-	-	-	-
	57	1287	15.2	14.5	-	-	-	-	-
	58	1063	CTD	CTD	-	-	-	-	-
	59	1028	15.7	9.8	-	-	-	-	-
	60	835	16.1	13.4	-	-	-	+	-
	Mean	1041	15.3	13.9					
	SD	145.6	0.58	2.45					
	n	10	9	9					
	Method †	2.7	2.7	2.7					
	P	0.426	0.899	0.103					

† See statistical analysis section in Experimental Procedure for explanation

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Hct L/L	Hb g/dL	RBC $x10^{12}$/L	MCH pg	MCHC g/dL	MCV fL
2F	41	0.473	16.0	8.06	19.9	33.9	58.6
	42	0.425	14.6	7.66	19.0	34.2	55.5
	43	0.468	16.0	8.49	18.9	34.2	55.1
	44	0.443	14.9	7.84	19.0	33.6	56.5
	45	0.445	15.4	7.84	19.7	34.7	56.8
	46	0.470	15.9	8.33	19.1	33.9	56.5
	47	0.422	14.5	7.53	19.2	34.3	56.1
	48	0.456	15.6	8.03	19.4	34.2	56.8
	49	0.446	15.0	7.89	19.0	33.7	56.5
	50	0.449	15.2	7.77	19.5	33.7	57.9
	Mean	0.450	15.3	7.94	19.3	34.0	56.6
	SD	0.0176	0.56	0.293	0.34	0.34	1.03
	n	10	10	10	10	10	10
	P						

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	WBC $x10^9$/L	N $x10^9$/L	L $x10^9$/L	E $x10^9$/L	B $x10^9$/L	M $x10^9$/L	LUC $x10^9$/L
2F	41	11.22	2.42	8.11	0.24	0.04	0.39	0.02
	42	5.97	0.45	5.34	0.08	0.00	0.08	0.02
	43	15.87	4.48	10.48	0.15	0.06	0.67	0.02
	44	8.98	0.97	7.73	0.14	0.02	0.09	0.02
	45	6.57	0.73	5.56	0.09	0.01	0.16	0.02
	46	11.85	2.39	8.87	0.13	0.03	0.41	0.02
	47	10.42	0.77	9.31	0.13	0.03	0.16	0.03
	48	4.90	1.08	3.52	0.07	0.00	0.21	0.01
	49	8.83	0.96	7.59	0.10	0.02	0.12	0.04
	50	10.58	0.64	9.61	0.10	0.05	0.12	0.06
	Mean	9.52	1.49	7.61	0.12	0.03	0.24	0.03
	SD	3.232	1.258	2.187	0.049	0.020	0.191	0.014
	n	10	10	10	10	10	10	10
	P							

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Plt x10^9/L	PT sec	APTT sec	Aniso-cytosis	Micro-cytosis	Macro-cytosis	Hypo-chromasia	Hyper-chromasia
2F	41	1191	15.1	11.5	-	-	-	-	-
	42	1038	15.2	13.7	-	-	-	-	-
	43	1178	15.9	18.1	-	-	-	-	-
	44	1135	16.0	11.9	-	-	-	-	-
	45	902	15.9	16.4	-	-	-	-	-
	46	974	15.5	14.8	-	-	-	-	-
	47	1036	15.5	10.6	-	-	-	-	-
	48	1031	15.7	17.4	-	-	-	-	-
	49	1195	15.1	14.4	-	-	-	-	-
	50	1038	14.8	14.5	-	-	-	-	-
	Mean	1072	15.5	14.3					
	SD	99.0	0.41	2.50					
	n	10	10	10					
	P								

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Hct L/L	Hb g/dL	RBC $x10^{12}$/L	MCH pg	MCHC g/dL	MCV fL
3F	61	0.447	15.4	7.88	19.5	34.4	56.8
	62	0.443	15.3	7.86	19.5	34.6	56.4
	63	0.447	15.2	7.87	19.3	33.9	56.8
	64	0.430	14.5	7.83	18.5	33.7	55.0
	65	0.489	16.4	8.51	19.3	33.6	57.4
	66	0.468	15.7	8.23	19.1	33.5	56.9
	67	0.439	15.2	8.06	18.8	34.6	54.5
	68	0.453	15.1	8.10	18.7	33.4	55.9
	69	0.433	15.0	7.69	19.6	34.7	56.3
	70	0.437	15.2	7.83	19.4	34.8	55.8
	Mean	0.449	15.3	7.99	19.2	34.1	56.2
	SD	0.0179	0.49	0.243	0.38	0.55	0.90
	n	10	10	10	10	10	10
	P						

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	WBC $x10^9$/L	N $x10^9$/L	L $x10^9$/L	E $x10^9$/L	B $x10^9$/L	M $x10^9$/L	LUC $x10^9$/L
3F	61	7.18	1.21	5.69	0.09	0.03	0.16	0.01
	62	7.44	0.50	6.69	0.11	0.01	0.11	0.02
	63	11.95	1.06	10.53	0.10	0.04	0.18	0.04
	64	7.34	0.41	6.72	0.10	0.02	0.07	0.02
	65	7.40	0.67	6.44	0.07	0.03	0.17	0.02
	66	7.51	0.32	6.96	0.06	0.02	0.10	0.05
	67	5.32	0.28	4.87	0.08	0.02	0.06	0.01
	68	6.67	0.39	6.06	0.10	0.01	0.10	0.02
	69	10.71	0.75	9.62	0.11	0.03	0.15	0.05
	70	10.43	0.84	9.31	0.09	0.03	0.12	0.04
	Mean	8.20	0.64	7.29	0.09	0.02	0.12	0.03
	SD	2.093	0.320	1.869	0.017	0.010	0.042	0.015
	n	10	10	10	10	10	10	10
	P							

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Plt x10^9/L	PT sec	APTT sec	Aniso-cytosis	Micro-cytosis	Macro-cytosis	Hypo-chromasia	Hyper-chromasia
3F	61	1116	15.5	16.0	-	-	-	-	-
	62	1133	14.7	17.4	-	-	-	-	-
	63	1111	15.0	16.2	-	-	-	-	-
	64	1318	15.2	14.2	-	-	-	-	-
	65	1121	15.1	13.3	-	-	-	-	-
	66	1087	15.5	16.9	-	-	-	-	-
	67	1099	15.8	17.0	-	-	-	-	-
	68	1115	15.2	14.7	-	-	-	-	-
	69	1075	15.7	16.9	-	-	-	-	-
	70	1009	16.2	15.2	-	-	-	-	-
	Mean	1118	15.4	15.8					
	SD	78.4	0.44	1.37					
	n	10	10	10					
	P								

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Hct L/L	Hb g/dL	RBC $x10^{12}$/L	MCH pg	MCHC g/dL	MCV fL
4F	71	0.436	15.0	7.72	19.4	34.5	56.4
	72	0.452	15.4	7.85	19.6	34.1	57.6
	73	0.435	14.7	7.83	18.8	33.9	55.5
	74	0.451	15.3	8.16	18.7	33.9	55.3
	75	0.427	14.4	7.42	19.4	33.7	57.6
	76	0.442	15.1	7.57	19.9	34.1	58.3
	77	0.439	15.3	7.64	20.1	34.9	57.5
	78	0.454	15.4	8.12	19.0	33.9	56.0
	79	0.443	15.0	7.96	18.9	34.0	55.6
	80	0.434	15.0	7.89	19.0	34.5	55.0
	Mean	0.441	15.1	7.82	19.3	34.2	56.5
	SD	0.0088	0.32	0.235	0.48	0.37	1.17
	n	10	10	10	10	10	10
	P						

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	WBC $x10^9$/L	N $x10^9$/L	L $x10^9$/L	E $x10^9$/L	B $x10^9$/L	M $x10^9$/L	LUC $x10^9$/L
4F	71	4.80	0.44	4.22	0.05	0.01	0.07	0.01
	72	8.22	1.36	6.56	0.09	0.02	0.15	0.03
	73	14.49	3.16	10.70	0.16	0.08	0.35	0.04
	74	3.92	0.71	3.03	0.09	0.01	0.08	0.00
	75	6.35	0.64	5.45	0.10	0.01	0.12	0.02
	76	11.30	0.53	10.45	0.08	0.02	0.16	0.06
	77	4.63	0.81	3.66	0.07	0.01	0.08	0.01
	78	5.99	0.84	4.84	0.16	0.02	0.11	0.02
	79	8.53	0.47	7.85	0.05	0.03	0.11	0.02
	80	4.41	0.46	3.81	0.05	0.01	0.07	0.01
	Mean	7.26	0.94	6.06	0.09	0.02	0.13	0.02
	SD	3.429	0.826	2.779	0.041	0.021	0.084	0.018
	n	10	10	10	10	10	10	10
	P							

APPENDIX 9 - continued

Haematology - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Plt x10^9/L	PT sec	APTT sec	Aniso-cytosis	Micro-cytosis	Macro-cytosis	Hypo-chromasia	Hyper-chromasia
4F	71	991	15.2	17.3	-	-	-	-	-
	72	1074	15.7	16.6	-	-	-	-	-
	73	1058	15.9	14.5	-	-	-	-	-
	74	1189	14.3	16.8	-	-	-	-	-
	75	1012	15.2	13.3	-	-	-	-	-
	76	1038	15.2	15.5	-	-	-	-	-
	77	1006	15.4	16.2	-	-	-	-	-
	78	1153	15.6	18.6	-	-	-	-	-
	79	1155	15.3	14.2	-	-	-	-	-
	80	1149	15.5	14.7	-	-	-	-	-
	Mean	1083	15.3	15.8					
	SD	72.9	0.43	1.62					
	n	10	10	10					
	P								

APPENDIX 10

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	ALP u/L	ALT u/L	AST u/L	gGT u/L	Bili µmol/L	Urea mmol/L	Creat µmol/L	Gluc mmol/L	Chol mmol/L	Trig mmol/L
1M	11	93	33	65	0	2	5.64	33	5.37	1.64	1.17
	12	69	35	58	0	1	4.26	26	7.37	2.25	0.75
	13	93	47	63	0	2	5.94	36	6.89	2.07	0.97
	14	88	47	82	1	2	4.32	30	7.12	1.48	0.88
	15	82	39	67	0	1	5.04	33	8.06	1.77	0.69
	16	76	40	69	0	2	5.23	35	5.32	1.85	1.31
	17	112	43	78	1	2	4.87	37	6.50	1.54	0.86
	18	105	51	69	0	2	6.59	30	5.51	1.36	0.38
	19	100	44	64	0	2	5.89	43	6.06	1.61	1.23
	20	103	43	67	1	2	6.64	40	6.57	1.48	0.55
	Mean	92	42	68	0	2	5.44	34	6.48	1.71	0.88
	SD	13.6	5.6	7.1	0.5	0.4	0.845	5.0	0.917	0.282	0.300
	n	10	10	10	10	10	10	10	10	10	10
	Method †	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7
	P	0.555	1.000	0.957	0.172	0.884	0.669	0.730	0.044	0.580	0.404

† See statistical analysis section in Experimental Procedure for explanation

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Na mmol/L	K mmol/L	Cl mmol/L	Ca mmol/L	Phos mmol/L	Total Prot g/L	Alb g/L	A/G Ratio
1M	11	141	5.6	97	2.45	2.14	72	39	1.18
	12	138	6.6	97	2.55	1.89	68	38	1.27
	13	137	6.5	98	2.50	2.29	68	38	1.27
	14	140	5.5	100	2.44	1.86	66	40	1.54
	15	139	6.0	99	2.41	2.32	65	37	1.32
	16	139	4.9	96	2.60	2.11	76	40	1.11
	17	139	5.6	102	2.39	2.00	67	37	1.23
	18	145	5.9	105	2.49	2.14	74	38	1.06
	19	142	4.6	101	2.61	1.95	71	40	1.29
	20	140	5.1	101	2.48	1.99	68	39	1.34
	Mean	140	5.6	100	2.49	2.07	70	39	1.26
	SD	2.3	0.65	2.8	0.075	0.158	3.6	1.2	0.133
	n	10	10	10	10	10	10	10	10
	Method †	2.7	2.7	2.7	2.7	2.7	2.7	8	2.7
	P	0.741	0.564	0.862	0.111	0.184	0.162	0.009	0.602

† See statistical analysis section in Experimental Procedure for explanation

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	ALP u/L	ALT u/L	AST u/L	gGT u/L	Bili µmol/L	Urea mmol/L	Creat µmol/L	Gluc mmol/L	Chol mmol/L	Trig mmol/L
2M	31	83	42	64	0	1	4.98	38	6.29	1.75	1.31
	32	73	37	54	1	2	4.57	33	7.39	1.78	2.17
	33	94	43	63	0	2	4.14	29	6.53	2.19	0.86
	34	96	52	90	0	2	5.76	33	7.33	1.37	1.11
	35	104	34	59	1	3	6.53	33	9.27	1.75	0.91
	36	65	38	55	1	3	4.81	30	6.73	2.31	0.70
	37	75	50	87	1	2	5.80	30	7.21	1.58	1.03
	38	77	37	69	1	2	4.47	37	5.99	1.54	0.90
	39	132	50	60	1	2	6.87	31	8.97	2.01	0.86
	40	89	40	68	0	1	5.77	31	7.50	1.59	0.38
	Mean	89	42	67	1	2	5.37	33	7.32	1.79	1.02
	SD	19.3	6.3	12.4	0.5	0.7	0.914	3.0	1.074	0.299	0.472
	n	10	10	10	10	10	10	10	10	10	10
	P								0.034		

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Na mmol/L	K mmol/L	Cl mmol/L	Ca mmol/L	Phos mmol/L	Total Prot g/L	Alb g/L	A/G Ratio
2M	31	137	5.6	98	2.51	1.84	71	39	1.22
	32	140	4.8	98	2.59	2.05	69	39	1.30
	33	141	5.7	101	2.39	2.31	67	37	1.23
	34	140	5.7	100	2.69	2.07	76	41	1.17
	35	138	5.1	98	2.58	1.89	69	38	1.23
	36	142	4.9	100	2.48	1.83	68	37	1.19
	37	141	6.2	101	2.48	2.12	63	35	1.25
	38	141	4.7	100	2.60	2.26	73	40	1.21
	39	136	6.3	98	2.51	1.73	67	40	1.48
	40	137	6.6	101	2.43	1.88	61	35	1.35
	Mean	139	5.6	99	2.53	2.00	68	38	1.26
	SD	1.9	0.67	1.3	0.089	0.194	4.4	2.1	0.093
	n	10	10	10	10	10	10	10	10
	P							n.s	

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	ALP u/L	ALT u/L	AST u/L	gGT u/L	Bili µmol/L	Urea mmol/L	Creat µmol/L	Gluc mmol/L	Chol mmol/L	Trig mmol/L
3M	1	69	30	58	0	1	5.54	37	7.13	1.73	1.19
	2	86	39	83	1	2	5.72	36	8.34	1.71	0.62
	3	97	44	55	0	2	5.88	29	7.29	1.39	0.55
	4	85	46	70	0	2	5.83	34	6.89	1.41	1.13
	5	119	56	68	0	3	4.28	31	8.03	1.53	0.60
	6	83	45	70	0	2	6.27	35	7.67	1.55	0.70
	7	104	37	58	1	1	5.21	36	7.55	1.63	1.03
	8	131	42	69	0	2	5.44	34	7.90	1.56	0.60
	9	74	42	61	0	3	7.18	35	6.92	1.89	0.99
	Mean	94	42	66	0	2	5.71	34	7.52	1.60	0.82
	SD	20.6	7.1	8.7	0.4	0.7	0.784	2.6	0.508	0.159	0.257
	n	9	9	9	9	9	9	9	9	9	9
	P								0.012		

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Na mmol/L	K mmol/L	Cl mmol/L	Ca mmol/L	Phos mmol/L	Total Prot g/L	Alb g/L	A/G Ratio
3M	1	142	4.7	100	2.52	1.92	72	38	1.12
	2	140	5.1	100	2.43	1.77	66	37	1.28
	3	139	5.6	98	2.52	2.23	70	38	1.19
	4	142	4.4	101	2.50	1.88	66	38	1.36
	5	140	5.4	98	2.50	1.78	73	38	1.09
	6	141	4.7	101	2.42	1.92	66	37	1.28
	7	138	6.9	100	2.50	1.89	72	36	1.00
	8	142	4.8	100	2.50	1.81	66	37	1.28
	9	138	5.3	97	2.60	2.27	71	38	1.15
	Mean	140	5.2	99	2.50	1.94	69	37	1.19
	SD	1.6	0.74	1.3	0.053	0.184	3.1	0.7	0.115
	n	9	9	9	9	9	9	9	9
	P							n.s	

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	ALP u/L	ALT u/L	AST u/L	gGT u/L	Bili µmol/L	Urea mmol/L	Creat µmol/L	Gluc mmol/L	Chol mmol/L	Trig mmol/L
4M	21	82	42	70	1	2	5.55	31	7.20	1.69	0.85
	22	100	44	53	0	2	5.57	30	7.51	2.08	0.91
	23	84	45	79	1	3	4.88	31	5.79	1.56	0.62
	24	78	38	69	1	1	4.41	31	7.78	1.39	1.16
	25	69	40	67	0	1	6.89	34	7.33	1.26	0.47
	26	71	35	59	1	3	5.78	30	7.31	1.44	0.57
	27	94	39	69	1	2	5.13	35	7.61	1.35	0.79
	28	68	39	57	2	2	5.93	43	9.03	2.10	1.06
	29	109	45	69	0	2	5.99	42	7.11	2.25	0.92
	30	82	54	80	0	1	7.82	34	7.13	2.04	0.45
	Mean	84	42	67	1	2	5.80	34	7.38	1.72	0.78
	SD	13.6	5.3	8.8	0.7	0.7	0.981	4.8	0.793	0.368	0.245
	n	10	10	10	10	10	10	10	10	10	10
	P								0.024		

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Na mmol/L	K mmol/L	Cl mmol/L	Ca mmol/L	Phos mmol/L	Total Prot g/L	Alb g/L	A/G Ratio
4M	21	142	4.9	103	2.44	1.74	63	37	1.42
	22	139	5.5	98	2.54	1.76	67	37	1.23
	23	142	4.7	101	2.36	2.02	60	36	1.50
	24	140	4.7	99	2.45	1.91	68	36	1.13
	25	139	5.2	100	2.49	1.98	69	37	1.16
	26	140	6.0	100	2.35	2.05	67	36	1.16
	27	140	5.3	101	2.37	1.80	65	36	1.24
	28	138	6.1	97	2.46	2.02	65	36	1.24
	29	141	6.2	102	2.47	2.05	69	39	1.30
	30	137	6.4	98	2.52	1.78	68	34	1.00
	Mean	140	5.5	100	2.45	1.91	66	36	1.24
	SD	1.8	0.64	1.9	0.066	0.129	2.9	1.3	0.144
	n	10	10	10	10	10	10	10	10
	P							<0.01	

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	ALP u/L	ALT u/L	AST u/L	gGT u/L	Bili µmol/L	Urea mmol/L	Creat µmol/L	Gluc mmol/L	Chol mmol/L	Trig mmol/L
1F	51	69	31	71	0	3	6.05	41	5.39	2.73	0.45
	52	56	40	61	1	2	7.05	44	5.74	2.04	0.66
	53	39	23	57	0	2	4.56	38	6.48	2.88	0.62
	54	45	36	69	1	2	5.30	36	6.33	2.37	0.65
	55	39	28	54	1	2	4.54	42	6.53	2.45	0.55
	56	40	53	149	0	3	6.08	41	5.40	2.59	0.70
	57	46	33	59	1	2	5.13	35	5.49	2.49	0.66
	58	35	26	59	0	3	5.41	38	6.17	1.99	0.71
	59	49	38	66	0	2	6.66	35	5.41	2.91	0.51
	60	41	36	56	1	2	5.63	43	7.12	1.87	1.95
	Mean	46	34	70	1	2	5.64	39	6.01	2.43	0.75
	SD	10.1	8.5	28.3	0.5	0.5	0.830	3.3	0.607	0.367	0.431
	n	10	10	10	10	10	10	10	10	10	10
	Method †	2.7	4.7	4.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7
	P	0.981	0.142	0.236	0.251	0.865	0.160	0.000	0.139	0.471	0.512

† See statistical analysis section in Experimental Procedure for explanation

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Na mmol/L	K mmol/L	Cl mmol/L	Ca mmol/L	Phos mmol/L	Total Prot g/L	Alb g/L	A/G Ratio
1F	51	139	4.5	97	2.59	1.93	76	47	1.62
	52	141	4.4	100	2.55	2.14	69	40	1.38
	53	138	4.1	98	2.69	1.86	78	45	1.36
	54	139	3.6	96	2.53	1.84	66	37	1.28
	55	140	4.1	100	2.62	1.80	79	44	1.26
	56	140	4.0	98	2.63	1.70	80	47	1.42
	57	137	4.2	100	2.66	1.82	78	40	1.05
	58	138	5.7	97	2.66	1.46	76	48	1.71
	59	141	4.8	101	2.55	2.08	67	38	1.31
	60	141	3.9	103	2.62	1.38	70	42	1.50
	Mean	139	4.3	99	2.61	1.80	74	43	1.39
	SD	1.3	0.59	2.2	0.054	0.240	5.3	4.0	0.189
	n	10	10	10	10	10	10	10	10
	Method †	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7
	P	0.902	0.134	0.764	0.019	0.114	0.399	0.437	0.601

† See statistical analysis section in Experimental Procedure for explanation

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	ALP u/L	ALT u/L	AST u/L	gGT u/L	Bili µmol/L	Urea mmol/L	Creat µmol/L	Gluc mmol/L	Chol mmol/L	Trig mmol/L
2F	41	47	47	69	1	2	5.18	41	5.09	2.27	1.28
	42	35	40	63	2	2	6.61	46	5.50	2.76	1.20
	43	49	39	65	1	2	6.40	43	5.09	1.83	0.49
	44	61	40	66	2	1	7.51	49	6.36	1.81	0.61
	45	47	34	63	1	3	6.29	42	6.74	2.45	0.96
	46	63	44	89	0	2	6.08	43	6.75	2.50	0.87
	47	33	33	82	1	3	5.48	43	6.42	2.63	0.91
	48	56	34	86	1	2	5.93	49	7.53	2.12	1.01
	49	33	136	132	1	3	5.99	43	7.73	2.46	1.32
	50	41	32	64	0	2	6.38	48	5.74	2.50	0.92
	Mean	47	48	78	1	2	6.19	45	6.30	2.33	0.96
	SD	11.1	31.3	21.5	0.7	0.6	0.637	3.0	0.934	0.322	0.268
	n	10	10	10	10	10	10	10	10	10	10
	P							0.007			

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Na mmol/L	K mmol/L	Cl mmol/L	Ca mmol/L	Phos mmol/L	Total Prot g/L	Alb g/L	A/G Ratio
2F	41	140	4.6	101	2.75	1.62	82	45	1.22
	42	138	5.9	100	2.69	1.60	76	45	1.45
	43	137	5.1	97	2.55	1.84	73	39	1.15
	44	137	4.5	98	2.50	1.96	68	39	1.34
	45	138	3.7	97	2.55	1.63	71	43	1.54
	46	143	5.0	103	2.61	1.96	74	41	1.24
	47	140	5.3	101	2.52	1.98	65	40	1.60
	48	142	5.9	102	2.55	1.61	75	41	1.21
	49	141	4.3	100	2.62	1.41	77	47	1.57
	50	138	4.9	99	2.59	1.61	73	41	1.28
	Mean	139	4.9	100	2.59	1.72	73	42	1.36
	SD	2.2	0.69	2.1	0.078	0.197	4.7	2.8	0.167
	n	10	10	10	10	10	10	10	10
	P				0.578				

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	ALP u/L	ALT u/L	AST u/L	gGT u/L	Bili µmol/L	Urea mmol/L	Creat µmol/L	Gluc mmol/L	Chol mmol/L	Trig mmol/L
3F	61	43	41	92	2	2	6.59	47	4.96	1.67	1.05
	62	47	29	66	1	2	5.23	44	6.06	2.92	1.11
	63	59	44	83	1	2	6.21	44	8.53	1.57	1.27
	64	48	43	77	2	3	6.51	45	6.53	1.73	0.39
	65	58	37	75	2	1	5.35	49	6.31	1.47	0.41
	66	58	31	78	0	3	7.44	46	6.17	2.83	0.88
	67	38	27	57	1	2	6.28	42	5.39	2.30	1.04
	68	42	28	66	1	2	5.10	45	6.83	2.82	1.16
	69	44	33	92	0	3	6.69	52	7.32	2.23	0.63
	70	42	38	75	0	1	6.69	53	8.07	1.84	0.62
	Mean	48	35	76	1	2	6.21	47	6.62	2.14	0.86
	SD	7.7	6.3	11.2	0.8	0.7	0.756	3.6	1.114	0.561	0.320
	n	10	10	10	10	10	10	10	10	10	10
	P							0.000			

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Na mmol/L	K mmol/L	Cl mmol/L	Ca mmol/L	Phos mmol/L	Total Prot g/L	Alb g/L	A/G Ratio
3F	61	137	5.6	96	2.57	1.67	79	44	1.26
	62	138	3.8	98	2.47	1.61	70	41	1.41
	63	141	4.3	101	2.52	1.63	74	44	1.47
	64	141	4.6	100	2.60	1.86	74	40	1.18
	65	140	4.4	101	2.56	1.93	64	39	1.56
	66	139	4.6	98	2.51	1.79	70	41	1.41
	67	138	5.5	100	2.62	1.73	70	40	1.33
	68	141	4.3	99	2.58	1.64	71	40	1.29
	69	139	6.3	99	2.45	1.67	71	40	1.29
	70	141	5.5	102	2.42	1.23	68	38	1.27
	Mean	140	4.9	99	2.53	1.68	71	41	1.35
	SD	1.5	0.78	1.8	0.067	0.189	4.0	1.9	0.114
	n	10	10	10	10	10	10	10	10
	P				0.012				

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	ALP u/L	ALT u/L	AST u/L	gGT u/L	Bili µmol/L	Urea mmol/L	Creat µmol/L	Gluc mmol/L	Chol mmol/L	Trig mmol/L
4F	71	39	34	63	1	1	7.12	57	5.85	3.09	1.32
	72	51	35	65	1	2	5.22	46	7.91	2.45	1.70
	73	75	40	77	1	2	8.22	51	7.44	1.60	0.45
	74	66	33	68	1	2	6.06	52	7.09	2.08	0.45
	75	43	44	72	0	2	7.15	46	7.15	2.60	2.19
	76	49	31	78	1	2	5.30	38	6.20	2.40	0.62
	77	44	34	78	1	2	5.90	43	5.82	2.25	0.59
	78	23	24	57	0	3	7.89	55	6.00	1.59	0.76
	79	32	27	61	1	3	6.84	48	7.99	2.29	0.76
	80	41	35	71	1	2	5.38	40	7.83	1.88	1.29
	Mean	46	34	69	1	2	6.51	48	6.93	2.22	1.01
	SD	15.2	5.7	7.5	0.4	0.6	1.090	6.2	0.883	0.461	0.589
	n	10	10	10	10	10	10	10	10	10	10
	P							0.000			

APPENDIX 10 - continued

Blood chemistry - individual values during Week 13 of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group / sex	Animal Number	Na mmol/L	K mmol/L	Cl mmol/L	Ca mmol/L	Phos mmol/L	Total Prot g/L	Alb g/L	A/G Ratio
4F	71	141	4.5	100	2.53	1.55	79	45	1.32
	72	140	4.6	102	2.60	1.26	78	45	1.36
	73	138	5.2	100	2.39	1.52	71	37	1.09
	74	138	5.0	100	2.48	1.39	73	41	1.28
	75	139	5.3	98	2.57	1.73	70	42	1.50
	76	140	4.9	99	2.61	2.06	76	43	1.30
	77	140	4.2	101	2.47	1.24	72	41	1.32
	78	140	5.3	99	2.59	1.89	77	42	1.20
	79	138	5.2	99	2.56	1.64	76	43	1.30
	80	143	4.4	98	2.51	1.16	70	40	1.33
	Mean	140	4.9	100	2.53	1.54	74	42	1.30
	SD	1.5	0.41	1.1	0.070	0.293	3.4	2.4	0.105
	n	10	10	10	10	10	10	10	10
	P				0.013				

Organ weights - individual absolute values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Terminal bodyweight	Adrenals	Brain	Epididymides	Heart	Kidneys	Liver	Spleen	Testes	Thymus
1M	11	497	0.056	2.34	1.381	1.429	3.02	15.60	0.767	3.92	0.114
	12	641	0.058	2.18	1.269	1.867	4.19	25.31	0.884	3.67	0.289
	13	561	0.048	2.29	1.177	1.495	2.95	20.14	0.601	3.30	0.343
	14	475	0.054	2.18	1.176	1.424	2.61	13.56	0.636	3.52	0.167
	15	561	0.043	2.15	1.387	1.742	2.98	18.52	0.805	3.49	0.341
	16	556	0.044	2.18	1.180	1.553	3.25	20.53	0.875	3.39	0.175
	17	439	0.048	2.07	1.212	1.326	2.29	13.40	0.791	3.40	0.281
	18	416	0.049	2.08	1.311	1.452	2.74	13.18	0.691	3.70	0.174
	19	497	0.049	2.23	1.241	1.339	2.80	15.68	0.875	3.72	0.211
	20	509	0.035	2.09	1.318	1.483	3.06	15.05	0.832	3.58	0.327
2M	31	569	0.061	2.13	1.166	1.880	3.37	18.55	0.904	3.22	0.290
	32	652	0.032	2.25	1.395	1.676	3.42	21.15	0.860	3.45	0.322
	33	472	0.042	2.06	1.184	1.468	2.81	15.40	0.736	3.95	0.216
	34	518	0.045	2.23	1.308	1.411	2.87	17.73	0.808	3.51	0.251
	35	526	0.067	2.10	1.202	1.490	3.09	18.86	0.793	3.77	0.190
	36	500	0.045	2.11	1.204	1.349	2.96	16.09	0.824	3.91	0.300
	37	555	0.040	2.00	0.973	1.613	3.08	18.61	1.050	3.17	0.247
	38	532	0.078	2.09	1.121	1.478	3.08	17.74	0.762	3.64	0.271
	39	491	0.044	2.04	1.408	1.433	3.27	18.85	0.637	3.30	0.156
	40	456	0.044	2.05	1.184	1.287	2.50	13.90	0.722	3.46	0.173

Organ weights - individual absolute values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Terminal bodyweight	Adrenals	Brain	Epididymides	Heart	Kidneys	Liver	Spleen	Testes	Thymus
3M	1	619	0.038	2.24	1.120	1.948	3.10	20.90	0.907	3.68	0.342
	2	444	0.045	2.05	1.273	1.245	2.79	13.52	0.486	3.39	0.115
	3	502	0.048	2.09	1.387	1.637	3.20	16.15	0.760	3.82	0.199
	4	465	0.046	2.09	1.200	1.466	2.90	16.33	0.693	3.30	0.163
	5	484	0.042	2.04	1.219	1.475	2.90	15.22	0.790	3.53	0.172
	6	507	0.048	2.13	1.292	1.333	2.85	15.86	0.929	3.81	0.173
	7	582	0.065	2.16	1.244	1.553	3.84	21.36	0.821	3.71	0.206
	8	494	0.039	1.91	1.223	1.613	2.99	14.89	0.657	3.26	0.227
	9	573	0.071	1.91	1.202	1.471	3.54	17.54	0.772	3.78	0.260
4M	21	446	0.047	2.05	1.304	1.404	2.63	12.76	0.593	3.47	0.166
	22	537	0.050	2.14	1.293	1.569	3.14	20.45	1.121	3.60	0.221
	23	410	0.043	2.08	1.062	1.381	2.20	11.72	0.600	3.30	0.194
	24	530	0.048	2.24	1.162	1.615	2.92	16.41	0.788	3.44	0.283
	25	534	0.056	2.07	1.271	1.662	3.28	17.29	0.646	3.79	0.192
	26	483	0.046	1.97	1.219	1.414	3.42	16.95	0.806	3.81	0.176
	27	549	0.045	2.22	1.312	1.561	2.94	16.00	0.916	3.58	0.215
	28	614	0.040	2.16	1.158	1.682	3.37	20.87	0.893	3.61	0.345
	29	459	0.043	2.01	1.366	1.449	2.57	13.50	0.554	3.82	0.276
	30	488	0.044	2.29	1.335	1.402	3.00	16.09	0.801	4.16	0.305

Organ weights - individual absolute values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Terminal bodyweight	Adrenals	Brain	Heart	Kidneys	Liver	Ovaries	Spleen	Thymus	Uterus + Cervix
1F	51	301	0.069	2.03	1.057	2.17	11.38	0.082	0.626	0.211	0.876
	52	299	0.073	1.97	0.946	2.17	10.43	0.100	0.464	0.215	0.473
	53	330	0.070	2.10	0.989	1.94	10.32	0.072	0.617	0.187	0.483
	54	326	0.061	2.00	1.121	2.17	11.22	0.083	0.562	0.331	0.989
	55	274	0.070	2.00	1.056	1.70	9.69	0.086	0.437	0.155	0.695
	56	302	0.052	2.08	1.010	1.93	9.40	0.087	0.518	0.172	0.559
	57	322	0.074	2.09	1.036	2.12	11.42	0.090	0.749	0.275	0.615
	58	272	0.066	2.07	0.957	1.94	9.84	0.077	0.598	0.303	0.927
	59	327	0.064	1.99	1.126	2.14	12.22	0.100	0.647	0.230	0.733
	60	295	0.058	1.87	0.959	1.62	9.09	0.092	0.505	0.246	0.845
2F	41	297	0.056	1.93	0.978	1.78	9.54	0.067	0.536	0.351	0.594
	42	303	0.061	2.07	1.090	1.86	11.40	0.087	0.457	0.169	0.689
	43	273	0.063	2.03	0.974	1.64	9.74	0.085	0.528	0.241	0.514
	44	279	0.053	1.89	1.066	1.73	9.02	0.089	0.535	0.310	0.829
	45	289	0.052	1.99	0.942	1.84	9.29	0.084	0.436	0.269	0.558
	46	296	0.076	2.08	0.953	1.90	11.19	0.088	0.637	0.295	0.888
	47	330	0.065	1.91	0.940	1.86	9.33	0.067	0.558	0.281	0.521
	48	297	0.063	2.02	1.016	2.03	9.65	0.091	0.611	0.257	0.571
	49	340	0.072	1.94	1.076	2.30	11.07	0.101	0.474	0.343	0.983
	50	318	0.065	2.01	1.097	1.68	10.62	0.088	0.700	0.221	0.628

Organ weights - individual absolute values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Group /Sex	Animal Number	Terminal bodyweight	Adrenals	Brain	Heart	Kidneys	Liver	Ovaries	Spleen	Thymus	Uterus + Cervix
3F	61	268	0.068	2.09	1.021	1.80	9.98	0.092	0.599	0.258	0.656
	62	266	0.050	1.92	0.847	1.54	8.61	0.066	0.495	0.278	0.750
	63	352	0.057	1.93	1.076	2.01	11.92	0.080	0.634	0.313	0.962
	64	307	0.077	2.08	1.062	2.23	10.20	0.124	0.629	0.421	0.639
	65	286	0.058	1.86	0.921	1.66	8.31	0.088	0.396	0.239	0.534
	66	314	0.071	1.97	0.962	2.02	9.70	0.094	0.638	0.182	0.488
	67	314	0.048	1.94	0.892	1.74	9.59	0.078	0.541	0.284	0.525
	68	276	0.050	1.97	0.847	1.69	8.20	0.074	0.462	0.225	0.543
	69	309	0.055	2.11	1.011	2.19	9.97	0.093	0.612	0.342	0.662
	70	316	0.060	2.17	1.049	2.00	9.99	0.080	0.647	0.293	1.092
4F	71	321	0.061	1.90	1.172	2.18	11.09	0.094	0.539	0.219	0.645
	72	363	0.066	2.00	1.136	2.16	10.98	0.070	0.587	0.391	0.713
	73	311	0.068	2.04	1.118	1.70	10.00	0.086	0.658	0.280	1.013
	74	267	0.054	1.86	0.934	1.74	8.45	0.123	0.492	0.171	0.472
	75	346	0.055	2.06	1.098	1.86	10.96	0.073	0.580	0.262	0.627
	76	313	0.075	2.10	1.096	2.26	11.35	0.091	0.512	0.249	0.601
	77	271	0.065	2.06	0.904	1.76	9.43	0.091	0.446	0.200	0.634
	78	280	0.056	2.06	0.923	1.64	9.09	0.092	0.409	0.097	0.400
	79	364	0.064	1.95	1.171	2.07	11.76	0.091	0.633	0.207	0.484
	80	314	0.067	2.07	0.956	1.96	9.35	0.089	0.641	0.232	0.910

APPENDIX 11 - continued

Organ weights - individual adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

GROUP	ANIMAL	TERMINAL BODY WT (g)	ADRENALS	BRAIN	EPIDIDYMID	HEART	KIDNEYS	LIVER	SPLEEN	TESTES	THYMUS
1M	11	496.7	0.056	2.34	1.381	1.431	3.02	15.65	0.768	3.918	0.11
	12	641.4	0.058	2.18	1.269	1.869	4.19	25.36	0.885	3.669	0.29
	13	560.7	0.048	2.29	1.177	1.497	2.95	20.19	0.602	3.303	0.34
	14	475.4	0.054	2.18	1.176	1.426	2.62	13.61	0.637	3.523	0.17
	15	560.7	0.043	2.15	1.387	1.744	2.99	18.57	0.806	3.488	0.34
	16	555.9	0.044	2.18	1.180	1.555	3.25	20.58	0.876	3.389	0.18
	17	438.7	0.048	2.07	1.212	1.328	2.29	13.45	0.792	3.401	0.28
	18	416.4	0.049	2.08	1.311	1.454	2.74	13.23	0.692	3.700	0.17
	19	497.0	0.049	2.23	1.241	1.341	2.80	15.73	0.876	3.720	0.21
	20	508.8	0.035	2.09	1.318	1.485	3.07	15.10	0.833	3.578	0.33
MEAN*		515.2	0.048	2.18	1.265	1.513	2.99	17.15	0.777	3.569	0.24
ST DEV*		66.36	0.0068	0.088	0.0813	0.1714	0.500	3.951	0.1017	0.1862	0.084
N		10	10	10	10	10	10	10	10	10	10
METHOD+			8	2.6	2.7	2.6	2.6	2.6	2.6	2.7	2.6
P			0.807	0.053	0.709	0.755	0.695	0.794	0.871	0.704	0.377

Adjusted values from ANCOVA using back-transformation where appropriate
* Mean and standard deviation adjusted to allow for the covariate.
\+ See Analysis Section in Experimental Procedure for explanation.

APPENDIX 11 - continued

Organ weights - individual adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

GROUP	ANIMAL	TERMINAL BODY WT (g)	ADRENALS	BRAIN	EPIDIDYMID	HEART	KIDNEYS	LIVER	SPLEEN	TESTES	THYMUS
2M	31	568.9	0.061	2.12	1.166	1.857	3.32	18.07	0.891	3.220	0.28
	32	652.3	0.032	2.24	1.395	1.653	3.36	20.67	0.847	3.452	0.31
	33	472.0	0.042	2.06	1.184	1.445	2.76	14.92	0.723	3.952	0.21
	34	517.7	0.045	2.22	1.308	1.388	2.82	17.26	0.795	3.512	0.24
	35	525.6	0.067	2.09	1.202	1.467	3.04	18.38	0.780	3.769	0.18
	36	499.9	0.045	2.10	1.204	1.326	2.91	15.61	0.811	3.905	0.29
	37	555.1	0.040	1.99	0.973	1.590	3.03	18.13	1.037	3.170	0.24
	38	532.1	0.078	2.08	1.121	1.455	3.03	17.26	0.749	3.639	0.26
	39	490.9	0.044	2.03	1.408	1.410	3.21	18.37	0.624	3.302	0.15
	40	455.7	0.044	2.04	1.184	1.264	2.45	13.42	0.709	3.457	0.17
MEAN*		527.0	0.050	2.10	1.215	1.486	2.99	17.21	0.796	3.538	0.23
ST DEV*		56.28	0.0142	0.081	0.1292	0.1731	0.276	2.066	0.1130	0.2742	0.056
N		10	10	10	10	10	10	10	10	10	10

Adjusted values from ANCOVA using back-transformation where appropriate
* Mean and standard deviation adjusted to allow for the covariate.

APPENDIX 11 - continued

Organ weights - individual adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

GROUP	ANIMAL	TERMINAL BODY WT (g)	ADRENALS	BRAIN	EPIDIDYMID	HEART	KIDNEYS	LIVER	SPLEEN	TESTES	THYMUS
3M	1	618.9	0.038	2.24	1.120	1.943	3.09	20.80	0.904	3.677	0.34
	2	443.6	0.045	2.04	1.273	1.240	2.77	13.41	0.483	3.393	0.11
	3	502.0	0.048	2.09	1.387	1.632	3.19	16.04	0.757	3.818	0.20
	4	464.6	0.046	2.09	1.200	1.461	2.89	16.23	0.690	3.296	0.16
	5	483.7	0.042	2.04	1.219	1.470	2.89	15.12	0.787	3.533	0.17
	6	506.5	0.048	2.13	1.292	1.328	2.83	15.76	0.926	3.809	0.17
	7	581.8	0.065	2.16	1.244	1.548	3.83	21.26	0.818	3.710	0.20
	8	494.2	0.039	1.91	1.223	1.608	2.98	14.79	0.654	3.261	0.23
	9	572.8	0.071	1.91	1.202	1.466	3.53	17.43	0.769	3.781	0.26
MEAN*		518.7	0.049	2.07	1.240	1.522	3.11	16.76	0.754	3.586	0.20
ST DEV*		58.92	0.0114	0.109	0.0738	0.2013	0.354	2.659	0.1346	0.2222	0.065
N		9	9	9	9	9	9	9	9	9	9

Adjusted values from ANCOVA using back-transformation where appropriate
* Mean and standard deviation adjusted to allow for the covariate.

APPENDIX 11 - continued

Organ weights - individual adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

GROUP	ANIMAL	TERMINAL BODY WT (g)	ADRENALS	BRAIN	EPIDIDYMID	HEART	KIDNEYS	LIVER	SPLEEN	TESTES	THYMUS
4M	21	445.9	0.047	2.06	1.304	1.429	2.69	13.28	0.608	3.471	0.17
	22	536.8	0.050	2.15	1.293	1.594	3.20	20.97	1.136	3.596	0.23
	23	409.6	0.043	2.09	1.062	1.406	2.25	12.24	0.615	3.299	0.20
	24	529.6	0.048	2.25	1.162	1.640	2.98	16.93	0.803	3.440	0.29
	25	533.5	0.056	2.08	1.271	1.687	3.34	17.81	0.661	3.793	0.20
	26	483.3	0.046	1.98	1.219	1.439	3.48	17.47	0.821	3.813	0.18
	27	548.8	0.045	2.23	1.312	1.586	3.00	16.52	0.931	3.577	0.22
	28	613.5	0.040	2.17	1.158	1.707	3.42	21.39	0.908	3.613	0.35
	29	458.5	0.043	2.02	1.366	1.474	2.63	14.02	0.569	3.822	0.28
	30	487.5	0.044	2.30	1.335	1.427	3.06	16.61	0.816	4.155	0.31
MEAN*		504.7	0.046	2.13	1.248	1.539	3.00	16.72	0.786	3.658	0.25
ST DEV*		59.30	0.0045	0.105	0.0956	0.1165	0.388	2.999	0.1778	0.2449	0.061
N		10	10	10	10	10	10	10	10	10	10

Adjusted values from ANCOVA using back-transformation where appropriate
* Mean and standard deviation adjusted to allow for the covariate.

APPENDIX 11 - continued

Organ weights - individual adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

GROUP	ANIMAL	TERMINAL BODY WT (g)	ADRENALS	BRAIN	HEART	KIDNEYS	LIVER	OVARIES	SPLEEN	THYMUS	UTERUS & C
1F	51	300.7	0.069	2.03	1.059	2.17	11.40	0.082	0.628	0.212	0.876
	52	299.0	0.073	1.97	0.948	2.17	10.46	0.100	0.466	0.216	0.473
	53	329.5	0.070	2.10	0.991	1.94	10.35	0.072	0.619	0.188	0.483
	54	325.9	0.061	2.00	1.123	2.18	11.25	0.083	0.564	0.332	0.989
	55	273.9	0.070	2.00	1.058	1.71	9.72	0.086	0.439	0.156	0.695
	56	301.5	0.052	2.08	1.012	1.93	9.43	0.087	0.520	0.173	0.559
	57	321.8	0.074	2.09	1.038	2.12	11.45	0.090	0.751	0.276	0.615
	58	272.0	0.066	2.07	0.959	1.95	9.87	0.077	0.600	0.304	0.927
	59	327.1	0.064	1.99	1.128	2.14	12.24	0.100	0.649	0.231	0.733
	60	295.3	0.058	1.87	0.961	1.62	9.11	0.092	0.507	0.247	0.845
MEAN*		304.7	0.066	2.02	1.028	1.99	10.53	0.087	0.574	0.233	0.720
ST DEV*		21.14	0.0070	0.070	0.0652	0.201	1.022	0.0091	0.0943	0.0569	0.1856
N		10	10	10	10	10	10	10	10	10	10
METHOD+			2.7	2.7	2.6	2.6	2.6	2.7	2.6	2.6	2.7
P			0.368	0.835	0.266	0.455	0.164	0.829	0.621	0.052	0.869

Adjusted values from ANCOVA using back-transformation where appropriate
* Mean and standard deviation adjusted to allow for the covariate.
\+ See Analysis Section in Experimental Procedure for explanation.

APPENDIX 11 - continued

Organ weights - individual adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

GROUP	ANIMAL	TERMINAL BODY WT (g)	ADRENALS	BRAIN	HEART	KIDNEYS	LIVER	OVARIES	SPLEEN	THYMUS	UTERUS & C
2F	41	297.0	0.056	1.93	0.985	1.79	9.64	0.067	0.542	0.354	0.594
	42	302.8	0.061	2.07	1.097	1.88	11.49	0.087	0.463	0.172	0.689
	43	272.9	0.063	2.03	0.981	1.65	9.84	0.085	0.534	0.244	0.514
	44	278.6	0.053	1.89	1.073	1.75	9.11	0.089	0.541	0.313	0.829
	45	289.3	0.052	1.99	0.949	1.86	9.38	0.084	0.442	0.272	0.558
	46	296.0	0.076	2.08	0.960	1.92	11.29	0.088	0.643	0.298	0.888
	47	329.9	0.065	1.91	0.947	1.88	9.43	0.067	0.564	0.284	0.521
	48	297.2	0.063	2.02	1.023	2.05	9.75	0.091	0.617	0.260	0.571
	49	339.8	0.072	1.94	1.083	2.32	11.17	0.101	0.480	0.346	0.983
	50	317.6	0.065	2.01	1.104	1.69	10.71	0.088	0.706	0.224	0.628
MEAN*		302.1	0.063	1.99	1.020	1.88	10.18	0.085	0.553	0.277	0.678
ST DEV*		21.32	0.0077	0.067	0.0637	0.193	0.891	0.0104	0.0831	0.0555	0.1657
N		10	10	10	10	10	10	10	10	10	10

Adjusted values from ANCOVA using back-transformation where appropriate
* Mean and standard deviation adjusted to allow for the covariate.

APPENDIX 11 - continued

Organ weights - individual adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

GROUP	ANIMAL	TERMINAL BODY WT (g)	ADRENALS	BRAIN	HEART	KIDNEYS	LIVER	OVARIES	SPLEEN	THYMUS	UTERUS & C
3F	61	267.6	0.068	2.09	1.031	1.82	10.11	0.092	0.607	0.263	0.656
	62	265.6	0.050	1.92	0.857	1.56	8.74	0.066	0.503	0.283	0.750
	63	352.4	0.057	1.93	1.086	2.04	12.05	0.080	0.642	0.318	0.962
	64	307.2	0.077	2.08	1.072	2.25	10.33	0.124	0.637	0.426	0.639
	65	286.4	0.058	1.86	0.931	1.69	8.44	0.088	0.404	0.244	0.534
	66	314.3	0.071	1.97	0.972	2.04	9.83	0.094	0.646	0.187	0.488
	67	314.1	0.048	1.94	0.902	1.77	9.73	0.078	0.549	0.289	0.525
	68	275.6	0.050	1.97	0.857	1.72	8.33	0.074	0.470	0.230	0.543
	69	308.6	0.055	2.11	1.021	2.21	10.10	0.093	0.620	0.347	0.662
	70	315.6	0.060	2.17	1.059	2.03	10.12	0.080	0.655	0.298	1.092
MEAN*		300.7	0.059	2.00	0.978	1.91	9.78	0.087	0.573	0.288	0.685
ST DEV*		26.93	0.0097	0.101	0.0876	0.234	1.094	0.0159	0.0875	0.0664	0.1993
N		10	10	10	10	10	10	10	10	10	10

Adjusted values from ANCOVA using back-transformation where appropriate
* Mean and standard deviation adjusted to allow for the covariate.

APPENDIX 11 - continued

Organ weights - individual adjusted values (g) for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

GROUP	ANIMAL	TERMINAL BODY WT (g)	ADRENALS	BRAIN	HEART	KIDNEYS	LIVER	OVARIES	SPLEEN	THYMUS	UTERUS & C
4F	71	321.2	0.061	1.90	1.154	2.14	10.83	0.094	0.523	0.210	0.645
	72	363.4	0.066	2.00	1.118	2.11	10.72	0.070	0.571	0.382	0.713
	73	311.4	0.068	2.04	1.100	1.65	9.74	0.086	0.642	0.271	1.013
	74	267.3	0.054	1.86	0.916	1.70	8.19	0.123	0.476	0.162	0.472
	75	346.1	0.055	2.06	1.080	1.81	10.71	0.073	0.564	0.253	0.627
	76	312.9	0.075	2.10	1.078	2.22	11.09	0.091	0.496	0.240	0.601
	77	271.3	0.065	2.06	0.886	1.72	9.18	0.091	0.430	0.191	0.634
	78	279.9	0.056	2.06	0.905	1.60	8.83	0.092	0.393	0.088	0.400
	79	363.6	0.064	1.95	1.153	2.02	11.50	0.091	0.617	0.198	0.484
	80	314.0	0.067	2.07	0.938	1.91	9.09	0.089	0.625	0.223	0.910
MEAN*		315.1	0.063	2.01	1.032	1.89	9.99	0.090	0.534	0.222	0.650
ST DEV*		35.19	0.0066	0.080	0.1084	0.224	1.123	0.0142	0.0848	0.0764	0.1910
N		10	10	10	10	10	10	10	10	10	10

Adjusted values from ANCOVA using back-transformation where appropriate
* Mean and standard deviation adjusted to allow for the covariate.

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 11 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 496.7 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Adrenals
Cortical Vacuolation, Minimal

Femur inc. Joint
Marrow - Fat Replacement, Minimal

Lungs + Bronchi
Arterial Mural Mineralisation, Present
Alveolar Macrophages, Focal, Minimal

Pituitary
Vacuolated Basophilic Cells, Minimal

Spleen
Extramedullary Haemopoiesis, Minimal

Urinary Bladder
Luminal Dilatation, Present

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 12 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 641.4 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Adrenals
Cortical Vacuolation, Minimal

Epididymides
Interstitial Inflammatory Cells, Multifocal, Minimal

Femur inc. Joint
Marrow - Fat Replacement, Minimal

Kidneys
Cortical Tubules with Hyaline Droplets, Minimal

Lungs + Bronchi
Arterial Mural Mineralisation, Present

Pancreas
Perivascular Inflammatory Cells, Focal, Minimal

Pituitary
Vacuolated Basophilic Cells, Minimal

Thyroids
Follicular Cell Hypertrophy, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 13 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 560.7 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Heart Myocardial Inflammatory Cells, Focal, Minimal
	LN Mandibular Plasmacytosis, Slight
Liver Swollen	Liver No significant lesions
	Pancreas Perivascular Inflammatory Cells, Focal, Minimal
	Prostate Inflammation, Focal, Minimal
	Thyroids Follicular Cell Hypertrophy, Minimal
	Urinary Bladder Luminal Dilatation, Present

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 14 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 475.4 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Epididymides
Interstitial Inflammatory Cells, Focal, Minimal

Femur inc. Joint
Marrow - Fat Replacement, Minimal

LN Mandibular
Plasmacytosis, Minimal

Pancreas
Perivascular Inflammatory Cells, Focal, Minimal

Spleen
Extramedullary Haemopoiesis, Minimal

Thyroids
Ectopic Thymic Tissue, Present

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 15 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 560.7 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	Heart Myocardial Inflammatory Cells, Focal, Minimal Kidneys Cortical Tubules with Hyaline Droplets, Minimal LN Mesenteric Sinus Erythrocytosis/Erythrophagocytosis, Minimal Thyroids Follicular Cell Hypertrophy, Minimal Urinary Bladder Luminal Dilatation, Present

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 16 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 555.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Femur inc. Joint Marrow - Fat Replacement, Minimal
	Heart Myocardial Inflammatory Cells, Focal, Minimal
	Kidneys Cortical Tubules with Hyaline Droplets, Minimal Cortical Tubular Basophilia, Focal, Minimal
LN Mandibular Enlarged, <13x10x5mm.	LN Mandibular Plasmacytosis, Slight
	Pituitary Vacuolated Basophilic Cells, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 17 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 438.7 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Epididymides Interstitial Inflammatory Cells, Multifocal, Minimal
	Kidneys Cortical Tubules with Hyaline Droplets, Minimal Cortical Tubular Basophilia, Focal, Minimal
	LN Mandibular Plasmacytosis, Moderate
	LN Mesenteric Mastocytosis, Minimal
Salivary Glands Oedematous	Salivary Glands No significant lesions
	Spleen Extramedullary Haemopoiesis, Minimal
	Thyroids Follicular Cell Hypertrophy, Minimal
	Urinary Bladder Luminal Dilatation, Present

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 17 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 438.7 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
continued from previous page...	
General Comments	General Comments
Moderate hairloss, Head	No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 18 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 416.4 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Femur inc. Joint Marrow - Fat Replacement, Minimal
LN Mandibular Enlarged, <14x10x4mm.	LN Mandibular Plasmacytosis, Moderate
	Lungs + Bronchi Alveolar Macrophages, Focal, Minimal
	Pituitary Vacuolated Basophilic Cells, Minimal
	Spleen Extramedullary Haemopoiesis, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 19 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 497.0 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Femur inc. Joint
Marrow - Fat Replacement, Minimal

LN Mandibular
Plasmacytosis, Slight
Mastocytosis, Minimal

Lungs + Bronchi
Arterial Mural Mineralisation, Present

Spleen
Extramedullary Haemopoiesis, Minimal

Thymus
Tissue missing

Thyroids
Follicular Cell Hypertrophy, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 20 Sex: Male Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 508.8 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Epididymides
Interstitial Inflammatory Cells, Focal, Minimal

Femur inc. Joint
Marrow - Fat Replacement, Minimal

Kidneys
Cortical Tubular Basophilia, Multifocal, Minimal

Prostate
Inflammation, Focal, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 31 Sex: Male Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 568.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
LN Mandibular	LN Mandibular
Enlarged, <14x7x4mm.	Plasmacytosis, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 32 Sex: Male Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 652.3 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 33 Sex: Male Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 472.0 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 34 Sex: Male Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 517.7 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 35 Sex: Male Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 525.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Adipose tissue	Adipose tissue
Strangulated Nodule(s), R. Epididymal, Yellow nodule, 7x4x2mm.	Fat Necrosis, Moderate

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 36 Sex: Male Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 499.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	Adrenals Cortical Vacuolation, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 37 Sex: Male Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 555.1 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Liver	Liver
Pale area(s), Left lobe, Area, 3x1mm.	Subcapsular Necrosis With Inflammation, Multifocal, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 38 Sex: Male Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 532.1 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
LN Mandibular Enlarged, <12x8x7mm.	LN Mandibular Plasmacytosis, Moderate
Thymus Dark area(s), Left lobe, Few punctate foci.	Thymus No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 39 Sex: Male Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 490.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	Adrenals
	Cortical Vacuolation, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 40 Sex: Male Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 455.7 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
General Comments	General Comments
Tail partially absent	No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 1 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 618.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Abdomen Contained fluid, Contains dark fluid and clotted blood.	Abdomen No significant lesions
LN Mandibular Enlarged, <16x8x3mm.	LN Mandibular Plasmacytosis, Moderate
Liver Ruptured, Caudate lobes, 2 poorly defined areas, on hilus.	Liver No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 2 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 443.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Liver Mass(es), Median lobe, Right, Firm mass, 5x4x3mm, on diaphragmatic surface. C/S parenchyma.	Liver No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 3 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 502.0 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 4 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 464.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Adrenals Cortical Vacuolation, Minimal
Thymus Dark area(s), Multiple punctate foci.	Thymus Haemorrhage, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 5 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 483.7 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Adrenals Cortical Vacuolation, Minimal
LN Mandibular Enlarged, <11x6x3mm.	LN Mandibular Plasmacytosis, Moderate
Thymus Dark area(s), Few punctate foci.	Thymus Haemorrhage, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 6 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 506.5 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	Adrenals Cortical Vacuolation, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 7 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 581.8 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Kidneys Enlarged, Right Pelvic dilatation, Right, Marked. Pale area(s), Right, Area, 3x2mm.	Kidneys Interstitial Inflammatory Cells, Focal, Minimal Hydronephrosis, Moderate Hydronephrosis-unilateral
LN Mandibular Enlarged, <15x9x4mm.	LN Mandibular Plasmacytosis, Slight
General Comments Fur stained, Upper dorsal thorax, Brown	General Comments No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 8 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 494.2 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Thymus Dark area(s), Few poorly defined areas.	Thymus Haemorrhage, Slight
General Comments Fur stained, Head, Brown Fur stained, Upper dorsal thorax, Brown	General Comments No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 9 Sex: Male Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 572.8 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	Adrenals Cortical Vacuolation, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 21 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 445.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Adrenals
Cortical Vacuolation, Minimal

Femur inc. Joint
Marrow - Fat Replacement, Minimal

Kidneys
Cortical Tubules with Hyaline Droplets, Minimal

LN Mandibular
Plasmacytosis, Slight

Lungs + Bronchi
Arterial Mural Mineralisation, Present

Pancreas
Acinar Replacement by Adipose Tissue, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 22 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 536.8 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Kidneys
Cortical Tubules with Hyaline Droplets, Minimal

LN Mandibular
Plasmacytosis, Slight

Liver
Arteritis/Periarteritis, Focal, Slight

Lungs + Bronchi
Arterial Mural Mineralisation, Present

Pancreas
Perivascular Inflammatory Cells, Focal, Minimal

Pituitary
Vacuolated Basophilic Cells, Minimal

Prostate
Inflammation, Focal, Minimal

Spleen
Extramedullary Haemopoiesis, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 23 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 409.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Adrenals
Cortical Vacuolation, Minimal

Femur inc. Joint
Marrow - Fat Replacement, Minimal

Heart
Epicardial Inflammatory Cells, Focal, Minimal

Kidneys
Cortical Tubular Basophilia, Focal, Minimal

LN Mandibular
Sinus Erythrocytosis/Erythrophagocytosis, Minimal

Pituitary
Developmental Cyst(s), Present

Spleen
Extramedullary Haemopoiesis, Minimal

Thyroids
Follicular Cell Hypertrophy, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 24 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 529.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Adrenals Cortical Vacuolation, Slight
	Femur inc. Joint Marrow - Fat Replacement, Minimal
	Heart Myocardial Inflammatory Cells, Focal, Minimal Myocardial Fibrosis, Focal, Minimal
LN Mandibular Enlarged, <13x9x5mm.	LN Mandibular Plasmacytosis, Moderate
	Lungs + Bronchi Arterial Mural Mineralisation, Present
	Pancreas Eosinophil Infiltration, Slight Oedema, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 25 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 533.5 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Femur inc. Joint
Marrow - Fat Replacement, Minimal

Kidneys
Cortical Tubules with Hyaline Droplets, Minimal
Cortical Tubular Basophilia, Focal, Minimal

LN Mandibular
Mastocytosis, Minimal

Pancreas
Perivascular Inflammatory Cells, Focal, Minimal

Spleen
Extramedullary Haemopoiesis, Minimal

Stomach
Mucosal Inflammation - Glandular Region, Focal, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 26 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 483.3 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Femur inc. Joint Marrow - Fat Replacement, Slight
	Heart Myocardial Inflammatory Cells, Focal, Minimal
Kidneys Pale area(s), Right, Area, 7x6mm, on mid-region.	Kidneys Cortical Tubules with Hyaline Droplets, Minimal
	Lungs + Bronchi Arterial Mural Mineralisation, Present
	Pancreas Periductal Inflammation / Fibrosis, Focal, Minimal Acinar Atrophy, Focal, Minimal
	Pituitary Vacuolated Basophilic Cells, Minimal
	Spleen Extramedullary Haemopoiesis, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 27 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 548.8 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Adrenals
Cortical Vacuolation, Minimal

Femur inc. Joint
Marrow - Fat Replacement, Minimal

Kidneys
Cortical Tubules with Hyaline Droplets, Minimal
Transitional Epithelial Inflammatory Cells, Minimal
Transitional Epithelial Hyperplasia, Minimal

Lungs + Bronchi
Arterial Mural Mineralisation, Present
Alveolar Macrophages, Focal, Minimal

Pancreas
Perivascular Inflammatory Cells, Focal, Minimal

Prostate
Inflammation, Focal, Minimal

Thyroids
Follicular Cell Hypertrophy, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 28 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 613.5 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Epididymides Interstitial Inflammatory Cells, Focal, Minimal
	Heart Myocardial Inflammatory Cells, Focal, Minimal Myocardial Fibrosis, Focal, Minimal
Kidneys Depression(s), Left, 2 punctate areas.	Kidneys Cortical Tubules with Hyaline Droplets, Minimal Cortical Tubular Basophilia, Focal, Minimal
	LN Mandibular Plasmacytosis, Minimal Sinus Erythrocytosis/Erythrophagocytosis, Minimal
	Lungs + Bronchi Arterial Mural Mineralisation, Present Alveolitis, Focal, Minimal Alveolar Epithelial Hyperplasia, Focal, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 28 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 613.5 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
continued from previous page...	
	LN Mesenteric Sinus Erythrocytosis/Erythrophagocytosis, Minimal
	Pancreas Perivascular Inflammatory Cells, Focal, Minimal
	Pituitary Vacuolated Basophilic Cells, Minimal
	Thyroids Follicular Cell Hypertrophy, Minimal
Thymus Dark area(s), Multiple punctate foci.	Thymus No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 29 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 458.5 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Adrenals
Cortical Vacuolation, Minimal

Femur inc. Joint
Marrow - Fat Replacement, Minimal

LN Mandibular
Plasmacytosis, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 30 Sex: Male Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 27-Feb-06 Day of Death: 92 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 487.5 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Epididymides Interstitial Inflammatory Cells, Focal, Minimal
	Heart Endocardial Inflammatory Cells, Focal, Minimal
	Kidneys Cortical Tubular Basophilia, Focal, Minimal
LN Mandibular Enlarged, <12x7x3mm.	LN Mandibular Plasmacytosis, Slight
	Lungs + Bronchi Arterial Mural Mineralisation, Present Alveolar Macrophages, Focal, Minimal
	Prostate Inflammation, Focal, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 51 Sex: Female Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 300.7 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Skin Scab(s), Head, Few punctate, dark areas.	Skin Scab(s), Focal, Minimal Epidermal Ulceration, Focal, Minimal Epidermal Hyperplasia, Focal, Minimal
Uterus Fluid distension, <4mm dia.	Uterus Luminal Dilatation, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 52 Sex: Female Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 299.0 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Kidneys Pelvic dilatation, Right, Moderate	Kidneys Hydronephrosis, Moderate Unilateral hydronephrosis Lungs + Bronchi Arterial Mural Mineralisation, Present Alveolitis, Focal, Minimal LN Mesenteric Sinus Erythrocytosis/Erythrophagocytosis, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 53 Sex: Female Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 329.5 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Kidneys Medullary Tubular Dilatation, Focal, Minimal
	Lungs + Bronchi Cholesterol Cleft Granuloma(ta), Focal, Minimal
Skin Scab(s), Head, Punctate focus, on left side.	Skin No significant lesions
General Comments Moderate hairloss, Head, Left	General Comments No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 54 Sex: Female Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 325.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Adrenals Pale Area(s), Multiple punctate foci.	Adrenals No significant lesions
	Heart Myocardial Inflammatory Cells, Focal, Minimal
Kidneys Pelvic dilatation, Right, Moderate	Kidneys Hydronephrosis, Slight Unilateral hydronephrosis
Uterus Fluid distension, <4mm dia.	Uterus Luminal Dilatation, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 55 Sex: Female Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 273.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Heart Myocardial Inflammatory Cells, Focal, Minimal
	LN Mandibular Plasmacytosis, Slight
	Pancreas Perivascular Inflammatory Cells, Focal, Minimal
	Spleen Extramedullary Haemopoiesis, Slight
General Comments Moderate hairloss, Head	General Comments No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 56 Sex: Female Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 301.5 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
General Comments	General Comments
Moderate hairloss, Head	No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 57 Sex: Female Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 321.8 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Kidneys Interstitial Inflammatory Cells, Focal, Minimal
LN Mandibular Enlarged, <14x8x4mm.	LN Mandibular Plasmacytosis, Slight
	Liver Inflammation, Portal, Focal, Minimal
Lungs + Bronchi Pale area(s), Left lobe, Multiple punctate foci. Pale area(s), Right caudal lobe	Lungs + Bronchi Alveolar Macrophages, Focal, Minimal
	Pancreas Perivascular Inflammatory Cells, Focal, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 58 Sex: Female Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 272.0 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	LN Mandibular Plasmacytosis, Minimal
	Spleen Extramedullary Haemopoiesis, Minimal
Uterus Fluid distension, <5mm dia.	Uterus Luminal Dilatation, Slight
General Comments Moderate hairloss, Head	General Comments No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 59 Sex: Female Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 327.1 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Femur inc. Joint Marrow - Fat Replacement, Minimal
	Kidneys Mineralisation, Corticomedullary, Focal, Minimal
General Comments Moderate hairloss, Head	General Comments No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 60 Sex: Female Dose Group: 1 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 295.3 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Adrenals Cortical Hypertrophy with Vacuolation, Focal, Minimal
	Femur inc. Joint Marrow - Fat Replacement, Minimal
	Spleen Extramedullary Haemopoiesis, Slight
Uterus Fluid distension, <5mm dia.	Uterus Luminal Dilatation, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 41 Sex: Female Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 297.0 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
LN Mandibular	LN Mandibular
Enlarged, <13x7x3mm.	Plasmacytosis, Moderate

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 42 Sex: Female Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 302.8 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
General Comments	General Comments
Moderate hairloss, Head	No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 43 Sex: Female Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 272.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
LN Mandibular	LN Mandibular
Enlarged, 11x9x5mm.	Plasmacytosis, Moderate

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 44 Sex: Female Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 278.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Uterus Fluid distension, <5mm dia.	Uterus Luminal Dilatation, Slight
General Comments Moderate hairloss, Head	General Comments No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 45 Sex: Female Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 289.3 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 46 Sex: Female Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 296.0 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
LN Mandibular	LN Mandibular
Enlarged, <11x7x3mm.	Plasmacytosis, Moderate
Uterus	Uterus
Fluid distension, <5mm dia.	Luminal Dilatation, Moderate

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 47 Sex: Female Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 329.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Kidneys Depression(s), Right, Punctate area.	Kidneys No significant lesions
Stomach Antrum white nodule(s), Punctate area.	Stomach No significant lesions
General Comments Moderate hairloss, Head, Right	General Comments No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 48 Sex: Female Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 297.2 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
LN Mandibular	LN Mandibular
Enlarged, <12x10x4mm.	Plasmacytosis, Moderate

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 49 Sex: Female Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 339.8 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Kidneys	Kidneys
Depression(s), 4 punctate areas.	Hydronephrosis, Moderate
Pelvic dilatation, Right, Moderate	Hydronephrosis-unilateral
Uterus	Uterus
Fluid distension, <4mm dia.	Luminal Dilatation, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 50 Sex: Female Dose Group: 2 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 317.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
LN Mandibular	LN Mandibular
Enlarged, <10x7x2mm.	Plasmacytosis, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 61 Sex: Female Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 267.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
LN Mandibular	LN Mandibular
Enlarged, <8x8x5mm.	Plasmacytosis, Moderate

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 62 Sex: Female Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 265.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 63 Sex: Female Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 352.4 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Stomach Antrum white nodule(s), 4 punctate areas.	Stomach Ectopic Nonglandular Epithelium in Glandular Region, Focal, Multifocal, Slight
Uterus Fluid distension, <4mm dia.	Uterus Luminal Dilatation, Moderate

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 64 Sex: Female Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 307.2 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 65 Sex: Female Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 286.4 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 66 Sex: Female Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 314.3 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 67 Sex: Female Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 314.1 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
General Comments	General Comments
Moderate hairloss, Head, Right	No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 68 Sex: Female Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 275.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 69 Sex: Female Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 308.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 70 Sex: Female Dose Group: 3 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 315.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Uterus Fluid distension, <5mm dia.	Uterus Luminal Dilatation, Moderate
General Comments Moderate hairloss, Head	General Comments No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 71 Sex: Female Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 321.2 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY

ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED

HISTOPATHOLOGY

Heart
Myocardial Inflammatory Cells, Focal, Minimal

Pancreas
Perivascular Inflammatory Cells, Focal, Minimal

Salivary Glands
Inflammation, Focal, Minimal

Spleen
Extramedullary Haemopoiesis, Minimal

Trachea
Subepithelial Lymphocytic Infiltration, Focal, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 72 Sex: Female Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 363.4 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Lungs + Bronchi Pale area(s), Left lobe, Few punctate foci.	Lungs + Bronchi No significant lesions Pancreas Perivascular Inflammatory Cells, Focal, Minimal Basophilic Acinar Cells, Focal, Slight Salivary Glands Inflammation, Focal, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 73 Sex: Female Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 311.4 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Lungs + Bronchi Arterial Mural Mineralisation, Present
	Spinal C. Lumb. Degenerate Fibres, Minimal
Uterus Fluid distension, <5mm dia.	Uterus Luminal Dilatation, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 74 Sex: Female Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 267.3 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	Pancreas Periductal Inflammation / Fibrosis, Focal, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 75 Sex: Female Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 346.1 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
ANIMAL HAS NO MACROSCOPIC FINDINGS RECORDED	***ANIMAL HAS NO MICROSCOPIC FINDINGS RECORDED***

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 76 Sex: Female Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 312.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
Lungs + Bronchi	Lungs + Bronchi
Pale area(s), Multiple punctate foci.	Alveolar Macrophages, Focal, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 77 Sex: Female Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 271.3 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Parathyroids Both missing
	LN Mandibular Reactive Histiocytosis, Minimal
	Pancreas Perivascular Inflammatory Cells, Focal, Minimal
Uterus Fluid distension, <4mm dia.	Uterus Luminal Dilatation, Slight

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 78 Sex: Female Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 279.9 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Heart Myocardial Inflammatory Cells, Focal, Minimal
	Lungs + Bronchi Alveolar Macrophages, Focal, Minimal
	Pancreas Perivascular Inflammatory Cells, Focal, Minimal
General Comments Moderate hairloss, Forelimb(s)	General Comments No significant lesions

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 79 Sex: Female Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 363.6 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Heart Myocardial Inflammatory Cells, Focal, Minimal
LN Mandibular Enlarged, <11x5x3mm.	LN Mandibular No significant lesions
	Pancreas Perivascular Inflammatory Cells, Focal, Minimal
	Pituitary Perivascular Inflammatory Cells, Focal, Minimal

Macropathology and histopathology - individual findings for animals killed after 13 weeks of treatment

Group	:	1	2	3	4
Compound	:	Control	T-AMG, PPY 24900	T-AMG, PPY 24900	T-AMG, PPY 24900
Dose (mL/kg/day)	:	0	1.0	3.3	10.0

Animal Number: 80 Sex: Female Dose Group: 4 Phase: Treatment Sacrifice Status: Final phase sacrifice
Date of Death: 28-Feb-06 Day of Death: 93 Week of Death: 14 Subgroup: 1 Terminal Bodyweight: 314.0 g

PATHOLOGY OBSERVATIONS

MACROPATHOLOGY	HISTOPATHOLOGY
	Heart Myocardial Inflammatory Cells, Focal, Minimal
	Kidneys Interstitial Inflammatory Cells, Focal, Minimal
	Lungs + Bronchi Arterial Mural Mineralisation, Present
	Spleen Extramedullary Haemopoiesis, Minimal
Uterus Fluid distension, <7mm dia.	Uterus Luminal Dilatation, Slight

ANNEX 1

Formulation Chemistry

Process Support Laboratories
Enzyme Analytical Laboratory

MCTo
2006-05-03
Luna no. 2006-18943-01

Huntingdon Life Sciences Study no: NVZ/0028
Novozymes reference no.: 20056035

Principal Investigator investigation Report

T-AMG, PPY 24900
Toxicity Study by
Oral Administration to CD rats for 13 weeks.

Analysis of samples returned from Huntingdon Life Sciences.

Content:

1 GLP COMPLIANCE

Study no. 20056035:

This investigation was conducted at the Process Support Laboratories, Enzyme Analytical Laboratory, Novozymes A/S, in compliance with OECD principles of Good Laboratory Practice, ENV/MC/CHEM (98) 17.

20060503
Date

2 Quality Assurance statement

QUALITY ASSURANCE STATEMENT

Report:	HLS Study no. NVZ/0028 T-AMG, PPY 24900, Toxicity study by oral administration to CD rats for 13 weeks. - Analysis of samples.
STUDY NUMBER	NVZ/0028
REFERENCE NUMBER	20056035

The conduct of this study has been subject to appropriate inspections and the report has been reviewed according to the relevant Standard Operation Procedures of Novozymes A/S Quality Assurance..

Inspection/Audit	Dates of inspection	Dates of Audit Report signed by Principle Scientist	Dates of Audit Report signed by Management
Analysis, Enzyme act.	22 March 2006	30 March 2006	30 March 2006
Report	2 May 2006	3 May 2006	3 May 2006

3. May 2006
Date

Quality Assurance

3 General information

Sponsor Monitor:
Nina Berg
Safety & Toxicology
Novozymes A/S
Krogshøjvej 36, 2880 Bagsværd
E-mail: NIB@novozymes.com

Principal Investigator:
Maria Camilla Tonsgaard
Process Support Laboratories
Enzyme Analytical Laboratory
Novozymes A/S
Krogshøjvej 36, 2880 Bagsværd
E-mail: mcto@novozymes.com

Study Director:
N. Hughes, H.N.C.
Huntingdon Life Sciences Ltd.
Woolley Road
Alconbury
Huntingdon
Cambridgeshire
PE28 4HS
ENGLAND
E-mail: HughesN@UKOrg.Huntingdon.com

Laboratory:
Enzyme Analytical Laboratory (EAL)
Process Support Laboratories
Novozymes A/S
Krogshøjvej 36, 2880 Bagsværd

Personel:
Laboratory Technician: Allan Lorck

Approved by: 20060503
Date

Principal Investigator

4 Purpose

The purpose of this investigation is to determine whether the enzyme activity (AGU/g) in the dose solutions from week 1, 6 and 13 are approximately equal and to check if the activity of the high 100% dose solution complies with the enzyme activity of the tox-batch.

5 Sample Handling

Sample description

During the study 36 samples were taken out for analysis of activity:
There are four groups:

- High activity (approx. 100%)
- Medium activity (approx. 33%)
- Low activity (approx. 10%)
- Control group (approx. 0%)

In week 1, 6, and 13, 3 samples of 10 ml are taken from each of the groups and labelled "1", "2" and "3".

Expected enzyme activities:

0% : 0 AGU/mL = 0 AGU/g
10%: 55.3 AGU/mL = 52.4 AGU/g
33%: 182.4 AGU/mL = 173 AGU/g
100%: 552.8 AGU/mL = 524 AGU/g

Sample transportation and registration

Samples from Huntingdon Life Sciences were received frozen at the Safety & Toxicology in Novozymes where the samples were registered.
The samples were stored frozen (-18°C) by Safety & Toxicology until transfer to EAL for analysis.

Storage of samples for analysis

After registration in Enzyme Analytical Laboratory the samples were stored frozen (-18°C) until analysis.
Note: There were some problems with freezer, described in part 7.

Sample defrost

The samples were defrosted at room temperature; 2006-03-22.

Date of analysis

Analysis of the samples was carried out 2006-03-22.

6 Method

The analysis is performed according to the valid version of the PSL-SM-0131.01-D Version 6.0, "Amyloglucosidase aktivitet Colorimetrisk på Konelab" (AGU) (= EB-SM-0131.02).

The analytical principle is described by 3 reaction steps:

Step 1 is an enzyme reaction:
Amyloglucosidase (AMG), EC 3.2.1.3 (exo-α-1,4-glucan-glucohydrolase), hydrolyzes maltose to form α-D-glucose. After incubation (37.0°C ± 1.0°C and 360 sek.), the reaction is stopped with NaOH.

Step 2 and 3 result in an endpoint reaction:
Glucose is phosphorylated by ATP, in a reaction catalyzed by hexokinase. The glucose-6-phosphate formed is oxidized to 6-phosphogluconate by glucose-6-phosphate dehydrogenase. In this same reaction an equimolar amount of NAD^+ is reduced to NADH with a resulting increase in absorbance at 340 nm.

The samples are analysed as 2 weighings at 1 standard curve as specified for GLP samples in current version of PSL-SP-0598.01-D.

7 Deviations

There were some problems with freezer in the Enzyme Analytical Laboratory and the samples were not frozen in a period of time, around 12 hours.
A deviation report is made to document the deviation. The Monitor and the Study Director were advised about the deviation. There has been taking proper corrective and preventive action.

8 Results and discussions

All raw data are given in table 4.

No activity above the detection limit was found for the control group for any of the samples. The activities are given as < 0.21 AGU/g in the table 4.

The data from high, medium and low dosing solution are evaluated according to current regulation, described in PSL-SP-0107.01-D, version 5.0.

The data are evaluated with 3 statistics tests. One test evaluate if the 3 groups, high, medium and low are almost equal for 1, 6 and 13 weeks. Another test evaluate if the high solution is equal to the tox batch. The mean activities of the 3 dosing solutions are estimated.
Results from the test are given in table 1, 2, 3 and 4.

Table 1. Analysis results of each sample for the dose grups high, medium and low, given in AGU/g. Expected activity was 524, 173 and 52 AGU/g.

Week	Sample No.	High	Medium	Low
1	1	525	173	50.9
	2	530	184	50.5
	3	542	174	51.1
6	1	536	173	51.2
	2	543	172	50.9
	3	556	175	51.1
13	1	525	173	52.9
	2	529	174	52.0
	3	532	175	51.1

Table 2. Approximate 95% confidence intervals for ratios between activity in week 6 and 13. Reference: week 1.

Group	Week	Lower Limit	Upper Limit	Is there significant difference?
High	6	0,97	1,08	No
	13	0,95	1,06	No
Medium	6	0,95	1,06	No
	13	0,94	1,05	No
Low	6	0,95	1,06	No
	13	0,98	1,10	No

Table 3. Mean activity per group for groups high, medium and low. Expected activity was 524, 173 and 52 AGU/g.

Group High	Group Medium	Group Low
535	175	51,3

The level for the samples for high, medium and low were found close to the intended activity.

Table 4. 95% confidence interval for ratio between mean of group high and Tox-batch (Group high/Tox-batch):

Analysis result for Tox-batch	Number of standard curves for Tox-batch (K_{Tox})	Number of weighings per standard curve for Tox-batch (N_{Tox})	Mean of group High	Lower Limit	Upper Limit	Is there significant difference?
524	3	2	528	0,95	1,07	No

The mean of the high solution is not significant different to the tox batch.

9 Conclusion

The measured concentration of the dosing solutions expressed in enzyme activity units was found close to the intended content of test material formulations for the high, medium and low dose groups. The dosing solutions are equal on the tested level on week 1, 2 and 13. There is no significant difference between the high dosing solution and the Tox-batch.

10 Archive

Original Investigation Plan, raw data or exact copies and Investigation Report are archived in QM Central Archive Novozymes A/S. A copy of this Investigation Report is distributed to Huntingdon Life Sciences for inclusion in the final report.

Table 5. Single data determination, week 1.

Week	Sample Group No	Conc	Expected Conc. AGU/g	Analyse No.	Results AGU/g (*)	Mean result AGU/g (*)
1	1	0%	0	23-34332	< 0.21	< 0.21
					< 0.21	
	2			23-34333	< 0.21	< 0.21
					< 0.21	
	3			23-34334	< 0.21	< 0.21
					< 0.21	
	1	10%	52.4	23-34335	51.0	50.9
					50.8	
	2			23-34336	50.7	50.5
					50.3	
	3			23-34337	51.0	51.1
					51.1	
	1	33%	173	23-34338	175	173
					172	
	2			23-34339	184	184
					184	
	3			23-34340	173	174
					175	
	1	100%	524	23-34341	524	525
					525	
	2			23-34342	527	530
					532	
	3			23-34343	545	542
					539	

(*) All results are given with 3 significant numbers and the mean results are calculated on the raw data.

Table 6. Single data determination, week 6.

Week	Sample Group No	Conc	Expected Conc. AGU/g	Analyse No.	Results AGU/g (*)	Mean result AGU/g (*)
6	1	0%	0	23-34344	< 0.21	< 0.21
					< 0.21	
	2			23-34345	< 0.21	< 0.21
					< 0.21	
	3			23-34346	< 0.21	< 0.21
					< 0.21	
	1	10%	52.4	23-34347	51.2	51.2
					51.2	
	2			23-34348	51.0	50.9
					50.7	
	3			23-34349	50.7	51.1
					51.4	
	1	33%	173	23-34350	174	173
					173	
	2			23-34351	172	172
					173	
	3			23-34352	175	175
					176	
	1	100%	524	23-34353	536	536
					535	
	2			23-34354	538	543
					547	
	3			23-34355	553	556
					558	

(*) All results are given with 3 significant numbers and the mean results are calculated on the raw data.

Table 7. Single data determination, week 13.

Week	Sample Group No	Conc	Expected Conc. AGU/g	Analyse No.	Results AGU/g (*)	Mean result AGU/g (*)
13	1	0%	0	23-37102	< 0.21	< 0.21
					< 0.21	
	2			23-37103	< 0.21	< 0.21
					< 0.21	
	3			23-37104	< 0.21	< 0.21
					< 0.21	
	1	10%	52.4	23-37105	53.0	52.9
					52.8	
	2			23-37106	51.9	52.0
					52.0	
	3			23-37107	50.8	51.1
					51.3	
	1	33%	173	23-37108	174	173
					172	
	2			23-37109	174	174
					173	
	3			23-37110	175	175
					175	
	1	100%	524	23-37111	523	525
					527	
	2			23-37112	528	529
					530	
	3			23-37113	529	532
					535	

(*) All results are given with 3 significant numbers and the mean results are calculated on the raw data.

ANNEX 2

Certificate of analysis

Safety & Toxicology

Date: 18. October 2005
Luna: 2005-40992-01
Ref.: KM

Documentation of Test Material

(Characterization Study no. 20058025)

Product:	**TOX BATCH**
Batch:	**PPY 24900**
Type of enzyme:	**T-AMG**
Host organism:	**Aspergillus niger**
Physical form / Colour:	**Brownish liquid at room temperature**

Activity:	**524 AGU/g**
Water (KF):	**85.6 % w/w**
Dry matter:	**14.4 % w/w**
Ash (600°C):	**0.5 % w/w**
Total Organic Solids (TOS):	**13.9 % w/w**
Specific gravity (g/ml):	**1.055 g/ml**
pH:	**6.1**
Total viable counts/g:	**<200**

Study Director

EYE RESEARCH CENTRE GLP COMPLIANCE STATEMENT 2005

THE DEPARTMENT OF HEALTH OF THE GOVERNMENT OF THE UNITED KINGDOM

GOOD LABORATORY PRACTICE

STATEMENT OF COMPLIANCE
IN ACCORDANCE WITH DIRECTIVE 2004/9/EC

LABORATORY	TEST TYPE
Huntingdon Life Sciences **Eye Research Centre** **Occold** **Eye** **Suffolk** **IP23 7PX**	**Analytical Chemistry** **Clinical Chemistry** **Ecosystems** **Environmental Fate** **Environmental Toxicity** **Mutagenicity** **Toxicology** **Phys/Chem Testing**

DATE OF INSPECTION

12th April 2005

A general inspection for compliance with the Principles of Good Laboratory Practice was carried out at the above laboratory as part of the UK GLP Compliance Programme.

At the time of inspection no deviations were found of sufficient magnitude to affect the validity of non-clinical studies performed at these facilities.

Head, UK GLP Monitoring Authority

HUNTINGDON RESEARCH CENTRE GLP COMPLIANCE STATEMENT 2005

THE DEPARTMENT OF HEALTH OF THE GOVERNMENT OF THE UNITED KINGDOM

GOOD LABORATORY PRACTICE

STATEMENT OF COMPLIANCE
IN ACCORDANCE WITH DIRECTIVE 2004/9/EC

LABORATORY	TEST TYPE
Huntingdon Life Sciences **Huntingdon Research Centre** **Woolley Road** **Alconbury** **Cambridgeshire** **PE28 4HS**	**Analytical Chemistry** **Clinical Chemistry** **Ecosystems** **Environmental Fate** **Environmental Toxicity** **Toxicology** **Phys/Chem Testing**

DATE OF INSPECTION

7th March 2005

A general inspection for compliance with the Principles of Good Laboratory Practice was carried out at the above laboratory as part of the UK GLP Compliance Programme.

At the time of inspection no deviations were found of sufficient magnitude to affect the validity of non-clinical studies performed at these facilities.

Head, UK GLP Monitoring Authority

Appendix 6
non-CCI version

Elements in Appendix 6 that are to be treated as confidential commercial information (CCI) are marked in highlighted text in the CCI version and the corresponding text as [redacted] in the non-CCI version.

Documentation regarding the production strain

1. Detailed description of the construction of the genetically modified production strain
2. Introduced DNA sequences in the production strain
3. DNA sequence of the *amgGT* gene and the amino acid sequence of the glucoamylase produced by *Aspergillus niger*, strain 41SaM2-54
4. Genetic stability of the production strain (Southern blot)

Appendix 6.1

Detailed description of the construction of the genetically modified production strain

6.1.1 The host organism

Taxonomy

The *Aspergillus niger* recipient strain, [redacted] was derived from strain BO-1, which was developed from the natural isolate *Aspergillus niger* C40-1. The BO-1 strain lineage has been used by Novozymes for more than 30 years and has given rise to a number of food enzyme production strains, which are used for production of previously evaluated and regulatory approved food enzymes. [redacted] has the following taxonomic characteristics:

Name: *Aspergillus niger*
Class: Eurotiomycetes
Order: Eurotiales
Genus: *Aspergillus*
Species: *niger*

The classification of *Aspergillus niger* BO-1 was confirmed by Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH [redacted].

Safety for humans and animals

The production species is non-pathogenic and has a long history of safe use for food. Furthermore, all genetic modifications are well characterized and specific, utilizing well-known genetic elements and plasmids. The introduced genetic material does not encode any known harmful or toxic substance. Hence, the genetically modified *Aspergillus niger* is considered a safe strain for enzyme production.

Furthermore, the GM strain is derived from a safe strain lineage (Pariza and Johnson., 2001) that includes the *Aspergillus niger* strain [redacted] Novozymes has used *Aspergillus niger* production strains derived from the BO-1 lineage for more than 30 years and has performed a number of safety studies on different enzyme products manufactured using *Aspergillus niger*. Table 1 below outlines some of the Novozymes products produced by *Aspergillus niger* strains and the safety studies conducted on those products.

Table 1 Novozymes products derived from *Aspergillus niger* strains where safety studies have been carried out. [a)] The predecessor strain shows strains in the GM construction that are in common with the BO-1 strain lineage, see Figure 1 below. [b)] At least the following: *in vitro* test for gene mutations in bacteria (Ames); *in vitro* test for chromosomal aberration or *in vitro* micronucleus assay; 13 week subchronic oral toxicity study in rats. The conclusions of these studies were in all cases favourable.

[redacted]

Figure 1 Overview of the *Aspergillus niger* safe strain lineage for strain [redacted] and different products made from strains in the lineage. The product number shown (Enzyme) corresponds to the number shown in Table 1. Only the strains that are in common with the strain [redacted] are shown. The number of GM steps between the depicted strains in the lineage is shown together with the genotype of each strain to the right.

Following the outline above, the BO-1 lineage is considered to be a safe strain lineage and therefore the use of food enzymes produced by strains from this lineage is safe. The fact that safety studies on different products manufactured from the same predecessor strain (e.g. strain BO-1,Figure 1) have been performed with favourable results strongly supports the safety of the *Aspergillus niger* lineage independent of the product manufactured. Similarly, the fact that safety studies have been performed on the same enzyme derived from different predecessor strains (e.g. glucoamylase derived from [redacted] emphasises that the controlled GM steps performed in the strain lineage do not cause safety issues.

For the enzyme preparation object of this application, toxicological testing of product produced by the current production strain was done and is reported in the application, Section A.5. All findings support the view that the production strain and its products are safe.

Genetic modifications

The recipient strain, [redacted] used in the construction of the *Aspergillus niger* production strain, was derived from the C40-1 parental strain through a combination of classical mutagenesis/selection and GM-steps. These steps included the inactivation of proteins involved in protein glycosylation and other strain improvements for product safety, purity, and stability.

[redacted]

[redacted]

6.1.2 Origin and donor of vector and inserts

The gene coding for the *Talaromyces emersonii* glucoamylase *amgGT* and the gene coding for the *Aspergillus niger* acid stable amylase asaA were introduced into the *Aspergillus niger* recipient strain [redacted]

Donor for the glucoamylase *amgGT* gene

The donor of the glucoamylase gene *amgGT* is a wild type *Talaromyces emersonii* strain (CBS 579.71). The enzyme gene is not protein engineered.

Donor for the promoter for *amgGT* expression

The promoter consists of the promoter from the *Aspergillus niger* strain BO-1 (CBS 120.49) *amyB* gene.

Donor for the transcriptional terminator for *amgGT*

The transcriptional terminator is derived from the *Aspergillus niger* strain BO-1 (CBS 120.49) *amgA* gene.

Donor for the acid stable amylase *asaA* gene

The donor of the acid stable amylase gene *asaA* is a wild type *Aspergillus niger* strain BO-1 (CBS 120.49). The enzyme gene is not protein engineered.

Donor for the promoter for *asaA* expression

The promoter consists of the acid stable amylase promoter from *Aspergillus niger* strain BO-1 (CBS 120.49).

Donor for the transcriptional terminator for *asaA*

The transcriptional terminator is derived from the *Aspergillus niger* strain BO-1 (CBS 120.49) *amgA* gene.

Vectors/insert

Vector pHUda81 is based on the *Escherichia coli* standard vector pUC19 (Vieira and Messing, 1987), containing an origin of replication (ORI) and an ampicillin resistance marker gene (amp). The pUC19 fragment including the ampicillin resistance gene was completely removed from the construct used for transformation of [redacted]

Vector pHUda211 is also based on the *Escherichia coli* standard vector puC19 (Vieira and Messing, 1987). However, only the ORI was used from the pUC19 vector and not the ampicillin resistance marker gene (amp). The ORI was completely removed from the construct used for transformation with pHUda211.

Selection of both expression plasmids in *Escherichia coli* was based on the URA3 (Rose et al., 1984) gene from *Saccharomyces cerevisiae* isolated from the general *Saccharomyces cerevisiae* cloning vector pYES2 (Invitrogen Corporation, Carlsbad, CA, USA) using an *Escherichia coli* strain with a *pyrF* mutation. The URA3 gene can complement the auxothrophic *pyrF* strain to uridine protothrophy.

The *oahA* gene deletion construct [redacted] is based on a commercial cloning vector from the general *Escherichia coli* phagemid vector pCR2.1 (Invitrogen Corporation, Carlsbad, CA, USA) that, in addition to the ORI, also harbours ampicillin and kanamycin resistance marker genes (amp and kan, respectively). The ORI, the ampicillin and the kanamycin resistance genes were completely removed from the construct used during the integration event of *pyrG* to disrupt the *oahA* gene.

Additionally, two marker genes were used in the different constructs, i.e. *amdS* and *pyrG*.

The *amdS* gene is from *Aspergillus nidulans* (Corrick et al., 1987) and codes for an acetamidase. It was isolated from the plasmid p3SR2, which was obtained from Prof. M. Hynes, Melbourne University. The *Aspergillus nidulans* acetamidase enzyme enables fungi to use acetamide as sole nitrogen and carbon sources (Kelly and Hynes, 1985). Acetamide, CH_3COONH_2, is hydrolysed by acetamidase to acetic acid and ammonia, which can be used as a carbon and nitrogen source, respectively.

The *pyrG* gene in [redacted] is from *Aspergillus niger* Bo-1.

The *pyrG* gene in pHUda211 is from *Aspergillus oryzae* IFO4117.

The *pyrG* gene encodes orotidine 5'-phosphate carboxylase and its deficient mutant is uridine auxotroph.

6.1.3 Introduced genetic sequence

The *amgGT* gene was cloned into the two plasmids pHUda81 and pHUda211 (Figure 2A and Figure 2B, respectively). The *asaA* gene was cloned into pHUda81 (Figure 2A). The individual genetic elements for plasmid pHUda81 and pHUda211 are listed in Table 3 and Table 4, respectively.

The inserted expression cassettes expressing the *amgGT* and *asaA* genes were integrated into the *Aspergillus niger* chromosomal DNA.

The two plasmids, pHUda81 and pHUda211, used to introduce the genes into the recipient strain are shown below.

6.1.4 Construction of the recombinant production organism

For the construction of the GM production strain 41SaM2-54, two transformation steps with expression cassettes for the two enzymes (glucoamylase and acid stable amylase) were performed. Additionally, disruption of the *oahA* gene encoding oxalic acid hydrolase was performed to eliminate the possibility for production of oxalic acid during fermentation. The steps from recipient strain [redacted] to GM production strain 41SaM2-54 are described below.

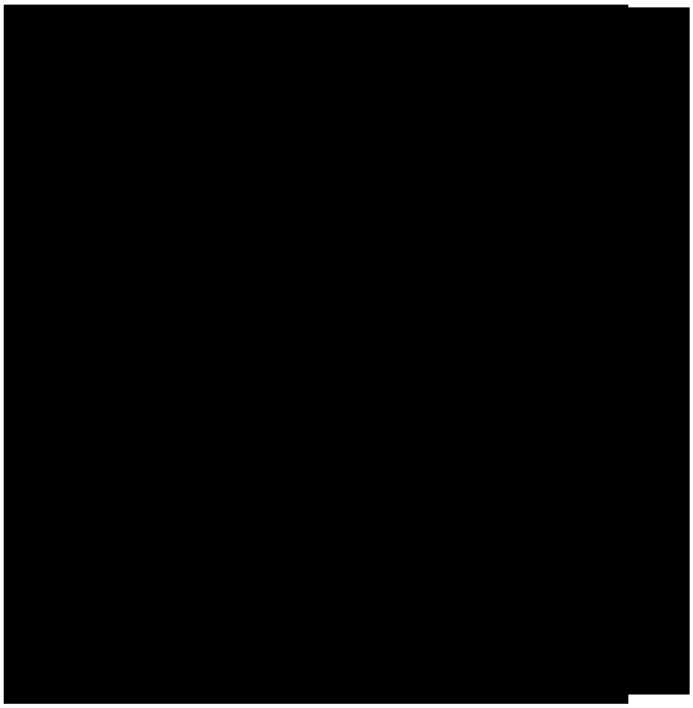


Figure 4 Strain construction lineage from *Aspergillus niger* recipient strain [redacted] to the GM strain 41SaM2-54. The DNA construction used in each genetic modification step or the selection method is shown to the right, whereas the step name is shown to the left.

Step 5 GM. Construction of strain [redacted] from [redacted]

The first recombinant DNA molecule (pHUda81, Figure 2A) was introduced into the recipient *Aspergillus niger* strain [redacted] by incubating protoplasts with the [redacted] plasmid pHUda81 (Figure 2A). As [redacted] lacks a functional *amdS* gene, it cannot grow with acetamide as the sole nitrogen source. Transformants are obtained upon the integration of multiple copies of the plasmid into the chromosomal DNA. Selection of transformants was therefore achieved by growing on a minimal medium with acetamide and subsequent screening for expression of glucoamylase. One transformant, strain [redacted] was selected. See under step 9 GM below for the demonstration of the introduction of both *asaA* and *amgGT* genes in the final GM strain.

Step 6. Isolation of strain [redacted] from [redacted]

Strain [redacted] was used to isolate a *pyrG*$^{-}$ strain using FOA selection, and strain [redacted] was isolated. Strain [redacted] displays the expected phenotype i.e., uridine requirement for growth.

Step 7 GM. Construction of strain [redacted] from [redacted]

Subsequently, strain [redacted] was transformed with a purified [redacted] fragment of [redacted] (Figure 3) containing the *oahA* gene disrupted with the *pyrG* gene, to inactivate the *oahA* gene coding for oxalic acid hydrolase gene. Transformants were selected in minimal medium without uridine. The transformants were screened for the lack of production of oxalic acid. Strain [redacted] was obtained where the *oahA* was disrupted [redacted]. Strain [redacted]

showed the expected phenotype (no production of oxalic acid) and was further characterized by Southern blot analysis.

The *oahA* probe used for Southern hybridization was prepared by PCR amplification using primers [redacted] and genomic DNA from strain BO-1 as DNA template. The [redacted] *oahA* probe is depicted in Figure 6A.

The *oahA* locus is characterized by a [redacted] hybridizing band using the *oahA* probe (Figure 5A). This band was observed for strain [redacted] as expected (Figure 6E, lane 1). Disruption of the *oahA* gene [redacted] from [redacted] (Figure 6B) with introduction of the *pyrG* gene should result in a change of the hybridizing band to [redacted] (Figure 6C and E).

The modification of the *oahA* locus in strain [redacted] corresponds to the disruption of the *oahA* gene with the introduction of *pyrG* at this locus, as shown by Southern analysis (Figure 6E, lane 2).

Step 8. Isolation of strain [redacted] from [redacted]
As above, strain [redacted] was used to isolate a *pyrG*$^{-}$ derivative by excision of the *pyrG* gene located at the disrupted *oahA* gene using FOA selection. Strain [redacted] was obtained and shown to require uridine for growth. Additionally, Southern blot was used to verify the excision of the *pyrG* gene from the modified *oahA* locus, resulting in the shift of the hybridizing band from [redacted] (Figure 6D and 6E, lane 3).

Step 9 GM. Construction of strain [redacted] from [redacted]
The isolated strain [redacted] was transformed with plasmid pHUda211 (Figure 2) using *pyrG* selection to introduce additional copies of the *amgGT* gene. A transformant, strain [redacted] was selected for high glucoamylase activity. As shown below (Figure 4), the final GM strain 41SaM2-54 contains the *amgGT* gene and additional copies of the *asaA* gene, as expected.

Step 10. Isolation of the GM strain 41SaM2-54 from [redacted]
Strain [redacted] was subjected to [redacted] screening for increased glucoamylase activity. The final GM production strain, 41SaM2-54 (Figure 4), which showed high expression of glucoamylase, was selected.

Genomic DNA was digested with [redacted] and hybridized to labelled *amgGT* specific gene probe (a [redacted] fragment obtained by PCR using pHUda81 as template DNA, Figure 2A). Strain 41SaM2-54 shows a single band corresponding to the *amgGT* gene ([redacted]), derived from both pHUda81 and pHUda211 (Figure 2). The size of the *amgGT* hybridizing band indicates that all integrated copies of the gene using pHUda81 and pHUda211 are maintained as full length (Figure 4, lane 2).

Furthermore, hybridization using the [redacted] *asaA* probe (Figure 2A) to [redacted] digested genomic DNA detected the endogenous gene present as a [redacted] band in both the recipient strain [redacted] and in the GM strain 41SaM2-54. A [redacted] band was also detected in strain 41SaM2-54, corresponding to the introduced copies of *asaA* using pHUda81 (Figure 4, lane 2).

6.1.5 Description of the production organism
The chromosome of the final production organism strain *Aspergillus niger* 41SaM2-54 has been modified [redacted] relative to the recipient strain. The modifications have led to the following changes in the strain:

- [redacted]
- [redacted]

The number of inserted copies for both *amgGT* and *asaA* was estimated by real-time qPCR. The insertion of the genes occurred randomly within the genome of the production strain. The presence of both *amgGT* and *asaA* was confirmed by Southern blot analysis (Appendix 6.2,Figure 5). A diagram of the *pyrG* expression cassette, which is integrated into the *oahA* locus in order to disrupt the wild type *oahA* gene in the *Aspergillus niger* chromosome, is given in Appendix 6.2 (Figure 6).

The DNA sequence of the *amgGT* gene and the amino acid sequence of the glucoamylase produced by *Aspergillus niger*, strain 41SaM2-54, is given in Appendix 6.3.

No antibiotic resistance genes were left in the final production strain as a result of the genetic modifications.

Identity and taxonomy of the production organism

The production strain is an *Aspergillus niger* carrying [redacted] of the *Talaromyces emersonii amgGT* gene coding for a glucoamylase and [redacted] of the *asaA* gene, including the endogenous copy present in the recipient strain.

Genetic stability and mobilisation and conjugation capability

The inserted recombinant DNA is genetically stable during fermentation, as the inserted DNA is integrated into the chromosome.

The genetic stability of the production strain was tested at large-scale fermentation. The strain stability during fermentation was analysed by Southern blot (Appendix 6.4). No instability of the strain was observed.

As all inserts are chromosomally integrated and lack a functional origin of replication, they cannot be transferred by conjugation to other organisms, nor can fragments replicate autonomously.

Antibiotic resistance gene

The resultant glucoamylase (*amgGT*) production strain 41SaM2-54 does not contain foreign antibiotic resistance genes. The absence of these genes in the production strain was verified by Southern blot analysis using the relevant antibiotic resistance gene probes.

References

Corrick CM, Twomey AP, Hynes MJ (1987) The nucleotide sequence of the *amdS* gene of *Aspergillus nidulans* and the molecular characterization of 5' mutations. Gene, 53, 63-71.

Kelly JM, Hynes MJ (1985) Transformation of *Aspergillus niger* by the *amdS* gene of *Aspergillus nidulans*. The EMBO Journal. 4 (2) 475-479.

Pariza MW and Johnson EA (2001) Evaluating the Safety of Microbial Enzyme Preparations Used in Food Processing: Update for a New Century. Regulatory Toxicology and Pharmacology, 33, 173-186.

Rose M, Grisafi P, Botstein D (1984) Structure and function of the yeast URA3 gene: expression in *Escherichia coli*. Gene, 29, 113-124.

Vieira J, Messing J (1987) Production of Single-Stranded Plasmid DNA. Methods in Enzymology, 153, 3-11

Appendix 6.4

Genetic stability of the production strain (Southern blot)

Summary

The genetic stability of the production strain was analysed by Southern blot of genomic DNA obtained from end of production samples, compared to a reference of genomic DNA from the production strain taken from the vial collection.

The Southern analysis of the end of production samples (EOP cells) and the reference sample showed no differences in the band pattern, thereby demonstrating the genetic stability of the inserted DNA in the 41SaM2-54 production strain.

Details

The genetic stability of the *Aspergillus niger* production strain 41SaM2-54 was analysed by Southern blot of genomic DNA obtained from reference cells and from EOP cells from three independent fermentation samples (batch 1, batch 2, batch 3).

Genomic DNA was isolated from culture suspension (i.e., end of production samples) that were allowed to sporulate and subsequently grow in liquid culture (Figure 7). This process adds additional generations to the cells used for the analysis allowing the analysis of genetic stability over the intended period of production.

Identical morphological characteristics were observed for all three EOP cells compared to the production strain reference (colony morphology on plates, sporulation and growth in liquid medium). The DNA derived from the EOP cells (Figure 7) was subsequently analysed by Southern blot analysis, comparing to DNA of the original production strain 41SaM2-54.

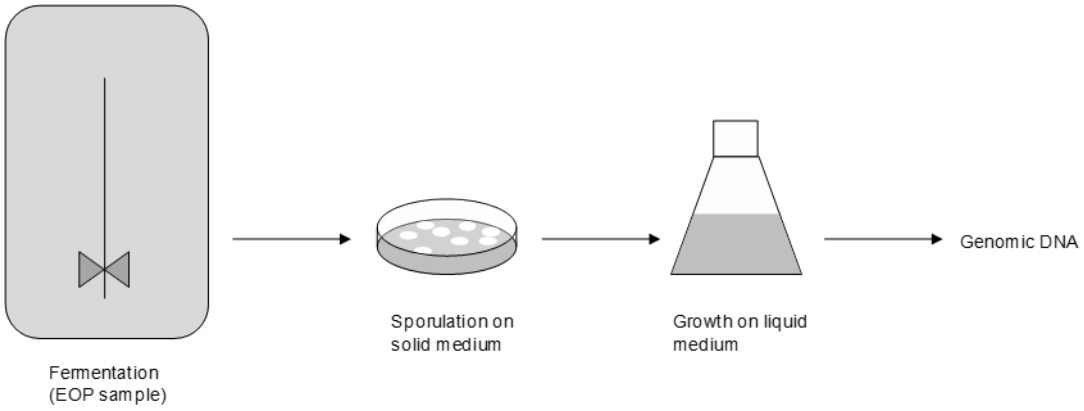


Figure 7 Overview of genomic DNA sample preparation for genetic stability analysis. A sample from end of production (EOP) from each pilot batch was taken and used to plate in solid medium to allow for growth of the strain and sporulation after 7 days. Spores were used to produce mycelium by growing in liquid medium 2-3 days. The mycelium was filtered and used for extraction of genomic DNA for Southern blot analysis.

End of production (EOP) samples from three independent batches were analysed (batch 1, batch 2, batch 3). Hybridisation to a gene-specific probe (*amgGT*, see Figure 8A) resulted in a number of bands derived from the number of copies of the *amgGT* gene inserted in 41SaM2-54 (Figure 8B).

Genomic DNA was purified from the samples and digested with [redacted]. DNA fragments were separated on an agarose gel and transferred to a Hybond-XL membrane by the Southern blotting method. Hybridization was performed using a [redacted] DNA fragment obtained by PCR

from pHUda211 covering the 5' end of the *amgGT* gene coding sequence (Figure 8) using a pair of primers [redacted] of pHUda211 (Figure 8A). After hybridization, the membrane was washed under stringent conditions and the blot was processed using a phosphoimager

Figure 8 (A) Map of plasmid pHUda81 and pHUda211 showing the fragment used as *amgGT* gene specific probe (only shown for pHUda211; a [redacted] PCR fragment) for Southern blot analyses. Only relevant restriction sites are shown. **(B)** Southern blot hybridization of genomic DNA digested with [redacted] and hybridized with the *amgGT* probe. Lane M: Molecular weight marker (λ BstE II); lane 1: batch 1; lane 2: batch 2; lane 3: batch 3; lane 4: strain 41SaM2-54. The sizes of the molecular marker are shown in kb to the left. The arrow to the right indicates the major expected band corresponding to the [redacted] fragment from both pHUda81 and pHUda211 (1.6 kb).

A major band observed in all DNA samples corresponds to the expected plasmid-derived band ([redacted]). The remaining bands observed are the result of non-homologous recombination that is characteristic of *Aspergillus niger* and other fungi.

Importantly, an identical hybridisation pattern of bands was obtained for the production strain and all three end of production samples analysed, demonstrating the stability of the inserted DNA (Figure 8B). This confirms that the procedure established for construction of production strains in *Aspergillus niger* results in the stable integration of the inserted DNA.

The Southern blot analysis of the end of production samples and the reference sample showed no differences in the band pattern (Figure 8B). Thus, the identity in the hybridisation pattern of the end of production samples derived from three independent batches and the production strain (41SaM2-54) demonstrates the genetic stability of the inserted DNA in the production strain 41SaM2-54.