An application to amend the Australia New Zealand Food Standards Code with a Glucose Oxidase preparation produced by a genetically modified *Trichoderma reesei*

AB Enzymes GmbH

Appendix 13 – CCI version

May 31, 2019

I. Table of contents

Elements in Appendix 14 that are to be treated as confidential commercial information (CCI) are marked in highlighted text in this CCI version (the corresponding text will be marked as ████████ in the non-CCI version).

II. The host organism

The recipient/host strain used for the genetic modification was ***Trichoderma reesei*** ██████████

a. Taxonomy

Trichoderma reesei is a hyper cellulolytic fungus which was found on deteriorating military fabrics such as tents and clothing. This original isolate, designated as QM6a, was initially named *Trichoderma viride*. Approximately 20 years later, QM6a was re-classified as *Trichoderma reesei*. In the 1980s, it was suggested that *Trichoderma reesei* should be placed into synonymy with *Trichoderma longibrachiatum* (Bissett 1991). Later however, evidence appeared that the two species were not identical (Meyer et al. 1992) and it was decided to go back to the *Trichoderma reesei* name. For a summary of *T. reesei´s* taxonomy, see Druzhinina et al. (2005).

Taxonomic studies have shown that the species *Trichoderma reesei* consists only of this single isolate QM6a and its derivatives (e.g. Rut Series, Montenecourt and Eveleigh, 1977, 1979; QM9123 and QM9414, Mandels *et al.*, 1971 – as reviewed by Nevalainen et al. (1994). The American Type Culture Collection (ATCC) designation for this original strain of *Trichoderma reesei* QM6a is ATCC 13631.

Taxonomic studies have shown that the species *Trichoderma reesei* consists only of a single isolate QM6a and its derivatives (Nevalainen et al. 1994).

The taxonomic classification of the *T. reesei* is: *Hypocreaceae, Hypocreales, Hypocreomycetidae, Sordariomycetes, Pezizomycotina, Ascomycota*, Fungi, according to Index *Fungorum* database.

Synonyms[1]: *Trichoderma reesei* is the species name given to the anamorphic form (the form which reproduces asexually) of the fungus whose teleomorphic form (the form which reproduces sexually) is now understood to be *Hypocrea jecorina*.

Trichoderma reesei was formerly known as *Trichoderma longibrachiatum*.

The *T. reesei* host strain is characterized by the Centraalbureau voor Schimmelcultures (CBS) in the Netherlands as *Trichoderma reesei*. It was identified based on the sequences of Internal Transcribed Spacer 1 and 2 and the 5.8S gene and Translation Elongation Factor 1α

b. Construction of the host

The *Trichoderma reesei* host strain is a mutant derived from *T. reesei* QM6a.

III. Origin and donor of vector and inserts

a. The enzyme gene

Origin

Glucose oxidase gene: The *Penicillium amagasakiense* glucose oxidase ▮▮▮▮▮▮▮▮▮▮ encoding sequence was designed and synthesized using the preferred codon usage for *Trichoderma reesei*.

Allergenicity

In order to specifically evaluate the risk that the xylanase enzyme will cross react with known allergens and induce a reaction in an already sensitized individual, sequence homology testing to known allergens was performed. This test used a 80 amino acid (aa) sliding window search as well as conventional FASTA (overall homology), with the threshold of 35% homology and scanning with 6 mer for exact matches as

---

[1] Reference: Mycobank taxonomic database (see:
 http://www.mycobank.org/Biolomics.aspx?Table=Mycobank&Page=200&ViewMode=Basic).

recommended in the most recent literature (Food and Agriculture Organization of the United Nations January/2001; Ladics et al. 2007; Goodman et al. 2008).

A sequence homology comparison test was then performed using a database of allergens from the Food Allergy Research and Resource Program (FARRP), University of Nebraska, Allergen Database (Version 16, January 27, 2016), which contains the amino acid sequences of known and putative allergenic proteins. No indication of an allergenic potential of the xylanase was detected (Appendix 13.1).

The resulting alignments of the full-length glucose oxidase protein to any allergenic proteins in the allergen database showed an identity of 30.7 % with Mala s 12 allergen precursors produced by the fungal species *Malassezia sympodialis*. Aalberse suggested "*cross-reactivity is rare below 50% amino acid identity and in most situations requires more than 70% identity*" (Aalberse 2000) making it unlikely that the glucose oxidase in question can be presumed to be allergenic based on full-length sequence relatedness to known allergens.

Using the 80-mer sliding window analysis the glucose oxidase protein sequence, one identity match of 41.2% with Mala s 12 allergen precursor produced by the fungal species *Malassezia sympodialis* was identified. As recommended by the FAO/WHO, a possible cross-reactivity has to be considered, when there is more than 35% identity in the amino acid sequence of the expressed protein using an 80 amino acids window and a suitable gap penalty (Food and Agriculture Organization of the United Nations January/2001). This recommendation was challenged however recently. According to Ladics et al. (2007) by comparing the predictive value of a full-length (conventional) FASTA search to the 80-mer analysis "*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*". This judgement on the predictive inferiority of the 80-mer (35% threshold) approach was supported recently by Goodman, Tetteh (2011) who suggested "*because the purpose of the bioinformatics search is to identify matches that may require further evaluation by IgE binding, full-length sequence evaluation or an increase in the threshold from 35% identity toward 50% for the 80 amino acid alignment should be considered*" (Goodman, Tetteh 2011). Using the latter

recommendation the glucose oxidase in question would be below threshold even using the 80-mer sliding window approach.

In addition, the glucose oxidase protein sequence showed no perfect match to any known allergen when searching for a straight stretch of eight amino acids that could serve as potential IgE binding sites.

In summary therefore the bioinformatics approach to estimate potential allergenicity based on relatedness to known allergens and taking into account the most recent scientific recommendations on the interpretation of such data leads us to conclude that the glucose oxidase produced by *Trichoderma reesei* RF11400 is of no concern.

Conclusion:

Based on the results obtained from the bioinformatics approach to estimate potential allergenicity on relatedness to known allergens and taking into account the most recent scientific recommendations on the interpretation of such data, and based on the fact that the enzyme is typically denatured during the food manufacturing process and that any residual enzyme still present in the final food will be subject to digestion in the gastro-intestinal system, it is not likely that the glucose oxidase produced by *Trichoderma reesei* RF11400 under evaluation will cause allergic reactions after ingestion of food containing the residues of these enzymes.

b. Vector

Trichoderma reesei strain RF11400 was constructed for production of *Penicillium amagasakiense* derived glucose oxidase by introducing the encoding gene into the genome of the *Trichoderma reesei* host strain. The glucose oxidase gene was cloned into the ███████ vector, resulting in the plasmid █████████ Plasmid map, full plasmid sequence, and table identifying each component are provided in Appendix 13.2.

c. Promoter

The Glucose oxidase gene is expressed under the strong ████████████ promoter.

IV. Introduced genetic sequence

Standard molecular biology methods were used in the construction of the expression plasmid. The expression cassette fragment used in fungal transformation does not contain any vector derived sequences as it is isolated from the expression plasmid by restriction digestion and purification from an agarose gel. It is free from any harmful sequences and contains the following genetic materials:

— Penicillium amagasakiense glucose oxidase gene: The glucose oxidase gene was synthesized using the preferred codon usage for *Trichoderma reesei*. The structural gene consists of an open reading frame of ██████ and codes for a protein of ███ amino acids, ████████████████ ████████████████ To construct the expression vector the gene is fused at its 5´-end to the ████████████████ promoter. This strong promoter is used to drive glucose oxidase expression to obtain high yields of glucose oxidase enzyme.

— **Terminator**: To ensure termination of transcription the native *T. reesei* cbhI terminator is used.

— **Aspergillus nidulans amdS gene**: the gene has been isolated from *Aspergilus nidulans* VH1-TRSX6 (Kelly, Hynes 1985). *Aspergillus nidulans* is closely related to *Aspergillus niger*, which is used in industrial production of food enzymes. The gene codes for an acetamidase that enables the strain to grow on acetamide as a sole nitrogen source. This characteristic has been used for selecting the transformants. The product of the *amdS* gene, acetamidase, can degrade acetamide and is not harmful or dangerous. The *amdS* marker gene has been widely used as a selection marker in fungal transformations without any disadvantage for more than 20 years.

Of the above genetic materials, the glucose oxidase gene and *Aspergillus nidulans amdS* gene are not naturally present in the host genome.

The DNA fragments that have been transformed to *T. reesei* host strain are well characterized, the sequences of the genes are known, and the fragments are free from any harmful sequences.

Diagram of the expression cassette and tables outlining the introduced genetic elements, as well as the full sequence of the DNA that as introduced as a result of the genetic modification are provided in Appendix 13.2.

V. Construction of the recombinant production organism

The transformation of *T. reesei* host strain with the expression cassettes was performed as described by Penttilä et al. (1987) with the modifications described in Karhunen et al. (1993).

Southern blot analyses were performed to the genome of the production strain RF11400. Results indicated that ████ copies of the expression cassettes were integrated in the genome of strain RF11400 (Appendix 13.3.).

VI. Description of the production organism

a. Identity and taxonomy

Genus: *Trichoderma*
Species: *Trichoderma reesei*
Subspecies (if appropriate):   not applicable
Generic name of the strain: RF11400

The taxonomic classification of the *T. reesei* is: *Hypocreaceae, Hypocreales, Hypocreomycetidae, Sordariomycetes, Pezizomycotina, Ascomycota*, Fungi, according to Index *Fungorum* database.

The transformed production strain containing the glucose oxidase gene is *Trichoderma reesei* strain RF11400 (RH33311) which is deposited in the "Centraalbureau voor Schimmelcultures" (CBS) in the Netherlands with the deposit number CBS138879.

Trichoderma reesei RF11400 was constructed for glucose oxidase production. The production strain differs from its recipient strain ████ in its high glucose oxidase production capacity due to expression of the glucose oxidase gene driven by ████ promoter (████ expression cassette)

integrated into the RF11400 genome. RF11400 secretes high amounts of glucose oxidase into its culture supernatant, resulting in high glucose oxidase activity in the cultivation broth.

b. Stability of the genetic traits in the GMM

T. reesei strains are widely used in biotechnological processes because of their known stability.

The production strain RF11400 is stable in terms of genetic traits. The genetic materials in the expression cassettes have been integrated as part of the genome and are as stable as any natural gene. The integrated genetic materials are not acting as mobilisable elements and they do not contain mobilisable elements.

Potential changes in the genome of the production strain could theoretically occur during the propagation in the fermentation process. Therefore, Southern blot analysis was performed after fermentation process of the RF11400 strain (see Appendix 13.4). The results revealed that the glucose oxidase gene stays genetically stable in *T. reesei* genome over necessary time that is needed for industrial fermentation process of the RF11400 production strain.

Additionally, the stability is also followed as equal production of the glucose in a number of fermentation batches performed for the RF11400. The activity measurements from parallel successful fermentations showed that the productivity of the RF11400 strain remains unchanged.

c. Mobilisation and transfer capacity

The inserted DNA does not include any mobile genetic element. Additionally, it should be highlighted that *T. reesei* genome lacks a significant repetitive DNA component and no extant functional transposable elements have been found in the genome (Kubicek et al. 2011; Martinez et al. 2008). This results to low risk of transfer of genetic material.

d. Presence of acquired antimicrobial resistance genes

The review article by Nevalainen et al. (1994) reveals that some species belonging to *Trichoderma* genus are able to secrete various types of antibiotics in laboratory cultures. However, strains of *T. reesei* used in

industrial applications are proven to be absent of antibiotic activities (Hjortkjaer et al. 1986; Coenen et al. 1995). The absence of antibiotic activities, according to the specifications recommended by JECFA was also confirmed.

Additionally, no genes have been introduced during the genetic construction that encode antimicrobial resistance.

VII. Information on any Significant Similarity between the Amino Acid Sequence of the Enzyme and that of Known Protein Toxins

A homology search was performed from the non-redundant protein sequences database using the BLAST-P (protein – protein BLAST) program, v. 2.2.30 (http://blast.ncbi.nlm.nih.gov/). The amino acid sequence of the gluocose oxidase (███ amino acids) was used as the query sequence in the searches.

BLAST-P is a basic local alignment search tool. By using this tool identities between two protein sequences can be found if the proteins contain similar sequence stretches (domains) even though the overall sequence homology between the sequences might be very low.

To evaluate the significance of reported matches, control searches were conducted using five different shuffled versions of the glucose oxidase amino acid sequence as query sequences. In these shuffled sequences the glucose oxidase amino acid sequence is randomly rearranged but the overall amino acid composition of the sequence is retained. The range of E values obtained for the shuffled sequences represents the background control incidence of random hits that could be expected basing on the amino acids contained within the glucose oxidase protein. The shuffled sequences were not expected to have significant homology to any other proteins in the database. In general, an E value approaching zero indicates that there is a very low probability that such a match would occur randomly by chance. Conversely, large E-value closer to 1, or even above 1, conveys that there is a relatively large probability that such a match would occur by chance. For study methods and results, see Appendix 13.5.

According to the results obtained from the searches performed it can be concluded that the glucose oxidase protein does not show significant homology to any protein sequence identified or known to be a toxin.

VIII. Appendices

IX. References

A. Publication bibliography

Coenen, T. M.; Schoenmakers, A. C.; Verhagen, H. (1995): Safety evaluation of beta-glucanase derived from Trichoderma reesei: summary of toxicological data. In *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* 33 (10), pp. 859–866.

Food and Agriculture Organization of the United Nations (January/2001): Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology,. Rome, Italy.

Goodman, Richard E.; Vieths, Stefan; Sampson, Hugh A.; Hill, David; Ebisawa, Motohiro; Taylor, Steve L.; van Ree, Ronald (2008): Allergenicity assessment of genetically modified crops--what makes sense? In *Nature biotechnology* 26 (1), pp. 73–81. DOI: 10.1038/nbt1343.

Hjortkjaer, R. K.; Bille-Hansen, V.; Hazelden, K. P.; McConville, M.; McGregor, D. B.; Cuthbert, J. A. et al. (1986): Safety evaluation of Celluclast, an acid cellulase derived from Trichoderma reesei. In *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* 24 (1), pp. 55-63.

Karhunen, T.; Mäntylä, A.; Nevalainen, K. M.; Suominen, P. L. (1993): High frequency one-step gene replacement in Trichoderma reesei. I. Endoglucanase I overproduction. In *Molecular & general genetics : MGG* 241 (5-6), pp. 515-522.

Kelly, J. M.; Hynes, M. J. (1985): Transformation of Aspergillus niger by the amdS gene of Aspergillus nidulans. In *The EMBO journal* 4 (2), pp. 475–479.

Kubicek, Christian P.; Herrera-Estrella, Alfredo; Seidl-Seiboth, Verena; Martinez, Diego A.; Druzhinina, Irina S.; Thon, Michael et al. (2011): Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of Trichoderma. In *Genome biology* 12 (4), pp. R40. DOI: 10.1186/gb-2011-12-4-r40.

Ladics, Gregory S.; Bannon, Gary A.; Silvanovich, Andre; Cressman, Robert F. (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. In *Molecular nutrition & food research* 51 (8), pp. 985–998. DOI: 10.1002/mnfr.200600231.

Martinez, Diego; Berka, Randy M.; Henrissat, Bernard; Saloheimo, Markku; Arvas, Mikko; Baker, Scott E. et al. (2008): Genome sequencing and analysis of the biomass-degrading fungus Trichoderma reesei (syn. Hypocrea jecorina). In *Nature biotechnology* 26 (5), pp. 553–560. DOI: 10.1038/nbt1403.

Nevalainen, H.; Suominen, P.; Taimisto, K. (1994): On the safety of Trichoderma reesei. In *Journal of biotechnology* 37 (3), pp. 193–200.

Paloheimo, Marja; Mäntylä, Arja; Kallio, Jarno; Puranen, Terhi; Suominen, Pirkko (2007): Increased production of xylanase by expression of a truncated version of the xyn11A gene from Nonomuraea flexuosa in Trichoderma reesei. In *Applied and environmental microbiology* 73 (10), pp. 3215-3224. DOI: 10.1128/AEM.02967-06.

Penttilä, M.; Nevalainen, H.; Rättö, M.; Salminen, E.; Knowles, J. (1987): A versatile transformation system for the cellulolytic filamentous fungus Trichoderma reesei. In *Gene* 61 (2), pp. 155-164.

Zhang, Z.; Wang, Y.; Ruan, J. (1998): Reclassification of Thermomonospora and Microtetraspora. In *International journal of systematic bacteriology* 48 Pt 2, pp. 411–422. DOI: 10.1099/00207713-48-2-411.