**(12) PATENT** **(11)** Application No. **AU 199933267 B2**
**(19) AUSTRALIAN PATENT OFFICE** **(10)** Patent No. **753475**

(54) Title
**Thermostable phytases in feed preparation and plant expression**

(51)$^7$ International Patent Classification(s)
**A23K 001/165** **C12N 009/16**
**A01H 005/00**

(21) Application No: **199933267** (22) Application Date: **1999.03.22**

(87) WIPO No: **WO99/48380**

(30) Priority Data

| (31) Number | (32) Date | (33) Country |
|---|---|---|
| **0407/98** | **1998.03.23** | **DK** |
| **PA 1998 00806** | **1998.06.19** | **DK** |
| **PA 1998 01176** | **1998.09.18** | **DK** |
| **PA 1999 00091** | **1999.01.22** | **DK** |
| **PA 1999 00093** | **1999.01.22** | **DK** |

(43) Publication Date : **1999.10.18**
(43) Publication Journal Date : **1999.12.09**
(44) Accepted Journal Date : **2002.10.17**

(71) Applicant(s)
**Novozymes A/S**

(72) Inventor(s)
**Svend Petersen**

(74) Agent/Attorney
**SPRUSON and FERGUSON,GPO Box 3898,SYDNEY NSW 2001**

(56) Related Art
**WO 97/16981**
**EP 619369**
**EP 682876**

77761/99

**PCT**

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

# INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

**(51) International Patent Classification** $^6$ **:**
**A23K 1/165, A01H 5/00, C12N 9/16**

**A1**

**(11) International Publication Number:** **WO 99/48380**

**(43) International Publication Date:** 30 September 1999 (30.09.99)

**(21) International Application Number:** PCT/DK99/00154

**(22) International Filing Date:** 22 March 1999 (22.03.99)

**(30) Priority Data:**

| | | |
|---|---|---|
| 0407/98 | 23 March 1998 (23.03.98) | DK |
| PA 1998 00806 | 19 June 1998 (19.06.98) | DK |
| PA 1998 01176 | 18 September 1998 (18.09.98) | DK |
| PA 1999 00091 | 22 January 1999 (22.01.99) | DK |
| PA 1999 00093 | 22 January 1999 (22.01.99) | DK |

SEC 113

Novozymes A/S, Krogshøjvej 36
**(71) Applicant:** ~~NOVO NORDISK A/S [DK/DK]; Corporate Patents, Novo Allé,~~ DK-2800 Bagsværd (DK).

**(72) Inventor:** PETERSEN, Svend; Novo Nordisk a/s, Novo Allé, DK-2880 Bagsværd (DK).
~~LEHMANN, Martin~~
~~258 Sayre Drive Princeton NJ 08540~~
~~US~~

92615
3/8/00

92615
2/2/00

**(81) Designated States:** AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published**
*With international search report.*

IP AUSTRALIA
18 OCT 1999
RECEIVED

**(54) Title:** THERMOSTABLE PHYTASES IN FEED PREPARATION AND PLANT EXPRESSION

**(57) Abstract**

The use of thermostable phytases in the preparation of animal feed, and the expression in plants of such phytases. For preparation of animal feed, a thermostable phytase is added before or during the agglomeration step. Preferred processes are pelleting, extrusion and expansion. A transgenic plant expressing a thermostable phytase may be used diretly in animal feed preparation.

# Thermostable phytases in feed preparation and plant expression

## Technical Field

This application relates to thermostable phytases, viz. their use in processes for the production of animal feed, and their expression in plants.

## Background art

WO 91/14782 describes transgenic tobacco and rapeseed plants expressing a phytase derived from Aspergillus ficuum NRRL 3135. The transgenic tobacco seeds are fed to broilers.

US 5,824,779 describes in standard fashion how to produce transgenic alfalfa expressing the same A. ficuum phytase, and the preparation of a phytase-containing concentrate which can be used per se as an animal feed supplement.

EP 0 556 883 B1 describes a method for preparing feed pellets based on an extrusion technique. The addition of temperature sensitive agents, one example of which is phytase, takes place after extrusion of the feed pellets, and the sensitive agents are loaded onto the pellets under reduced pressure.

As acknowledged in EP 0 556 883 B1 the loss of activity of heat-sensitive substances during feed preparation processes is a well-known problem. The above EP-patent proposes to solve this problem by adding these substances under reduced pressure subsequent to the extrusion process. This solution, however, requires a liquid form of the sensitive substance, as well as the installation of additional expensive process equipment.

The present invention provides an improved process for preparing animal feed, as well as improved phytase-expressing transgenic plants.

## Summary of the Invention

Herein disclosed is a process of preparing an animal feed, which process comprises an agglomeration of feed ingredients, wherein a thermostable phytase is added before or during the agglomeration.

Thus, according to an embodiment of the invention, there is provided a process for preparing an animal feed, wherein a phytase with a Tm of at least 75°C is added before or during an agglomeration of feed ingredients, the Tm being measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE.

According to another embodiment of the invention, there is provided the use of a phytase with a Tm of at least 75°C in a process for preparing an animal feed, wherein the phytase is added before or during an agglomeration of feed ingredients, and wherein the Tm is measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE.

According to another embodiment of the invention, there is provided the use of a phytase with a Tm of at least 75°C in the manufacture of a preparation for use in a process for preparing an animal feed in which the phytase preparation is added before or during an agglomeration of feed ingredients, and wherein the Tm is measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE.

According to another embodiment of the invention, there is provided an expression construct which comprises a DNA construct encoding a phytase with a Tm of at least 75°C, the Tm being measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE, said DNA construct being operably linked to regulatory sequences capable of mediating expression of said phytase encoding sequence in at least one part of a plant. Vectors comprising the expression constructs of the invention are also provided.

According to another embodiment of the invention, there is provided a method of preparing a transgenic plant capable of expressing a phytase with a Tm of at least 75°C, the Tm being measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE, said method comprising the steps of

(i) isolating a nucleotide sequence encoding the phytase;

(ii) inserting the nucleotide sequence of (i) in an expression construct capable of mediating the expression of the nucleotide sequence in a selected host plant; and

(iii) transforming the selected host plant with the expression construct.

According to another embodiment of the invention, there is provided method of preparing a transgenic plant cell capable of expressing a phytase with a Tm of at least 75°C, the Tm being measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE, said method comprising the steps of

(i) isolating a nucleotide sequence encoding the phytase;

(ii) inserting the nucleotide sequence of (i) in an expression construct capable of mediating the expression of the nucleotide sequence in a selected host plant cell; and

(iii) transforming the selected host plant cell with the expression construct. Transgenic plant cells prepared by this method are also provided. Also provided is a transgenic plant or part thereof which comprises a DNA-construct encoding a thermostable phytase. In a preferred aspect, the transgenic plants are prepared by the methods of the invention, or are regenerated from, or comprise a plurality of transgenic cells obtained by the methods of the invention.

Thus, according to another embodiment of the invention, there is provided a transgenic plant which comprises a DNA-construct encoding a phytase with a Tm of at least 75°C, the Tm being measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE.

Transgenic seed obtained from transgenic plants of the invention are also provided.

According to another embodiment of the invention, there is provided a method of producing a phytase with a Tm of at least 75°C, said method comprising culturing one or more transgenic cells of the invention, or growing one or more transgenic plants of the invention, under conditions conducive to expression of said phytase in said plant cell, plant, or transgenic plant part or seed of said plant. Phytases with a Tm of at least 75°C produced by this method, and their use in the processes of the invention are also provided.

The transgenic plant or part thereof, e. g. seeds or leaves, may be used in the feed preparation process of the invention, to thereby provide in a preferred embodiment - at the same time a nutrient (feed ingredient) and the feed additive phytase.

## Brief description of the Figures

In the detailed description of the invention below, reference is made to the drawings, of which

Fig. 1 is a differential scanning calorimetry (DSC) chart of consensus phytase-1 and consensus phytase-10 ;

Fig. 2 a DSC of consensus phytase-10-thermo-Q50T and consensus phytase-10-thermo-Q50T-K91A;

Fig. 3 a DSC of consensus phytase-1-thermo [8]-Q50T and consensus phytase-1-thermo [8]-Q50T-K91A;

Fig. 4 a DSC of the phytase from A. fumigatus ATCC 13073 and of its a-mutant; and

Fig. 5 shows the design of the consensus-phytase-1 amino acid sequence;

| | |
|---|---|
| Fig. 6 | an alignment and the basidiomycete consensus sequence of five Basidiomycete phytases; |
| Fig. 7 | the design of the consensus-phytase-10 amino acid sequence; |
| Fig. 8 | an alignment for the design of consensus-phytase-11 (all Basidiomycete phytases were used as independent sequences using an assigned vote weight of 0.2 for each Basidiomycete sequence; still further the amino acid sequence of A. niger T213 was used); |
| Fig. 9 | the DNA and amino acid sequence of consensus-phytase-1-thermo(8)-Q50T-K91A; |
| Fig. 10 | the DNA and amino acid sequence of Consensus-phytase-10-thermo(3)-Q50T-K91A; |
| Fig. 11 | the DNA and amino acid sequence of A. fumigatus ATCC 13073 α-mutant; and |
| Fig. 12 | the DNA and amino acid sequence of Consensus-phytase-7 which comprises the following mutations as compared to Consensus-phytase-1: S89D, S92G, A94K, D164S, P201S, G203A, G205S, H212P, G224A, D226T, E255T, D256E, V258T, P265S, Q292H, G300K, Y305H, A314T, S364G, M365I, A397S, S398A, G404A, and A405S. |

## Detailed description of the invention

In the present context a "feed" or an "animal feed" means any natural or artificial diet, meal or the like intended or suitable for being eaten, taken in, digested, by an animal. Food for human beings is included in the above definition of feed.

"Animals" include all animals, be it polygastric animals (ruminants); or monogastric animals such as human beings, poultry, swine and fish. Preferred animals are the mono-gastric animals, in particular pigs and broilers.

The concept of "feed ingredients" includes the raw materials from which a feed is to be, or is, produced; or the intended, or actual, component parts of a feed. Feed ingredients for non-human animals are usually, and preferably, selected from amongst the following non-exclusive list:

plant derived products

such as seeds, grains, leaves, roots, tubers, flowers, pods, husks - and they may take the form of flakes, cakes, grits, flour, and the like;

animal derived products

such as fish meal, milk powder, bone extract, meat extract, blood extract and the like;

additives

such as minerals, vitamins, aroma compounds, and feed enhancing enzymes.

Phytic acid or myo-inositol 1,2,3,4,5,6-hexakis dihydrogen phosphate (or for short myo-inositol hexakisphosphate) is the primary source of inositol and the primary storage form of phosphate in plant seeds and grains. In the seeds of legumes it accounts for about 70% of the phosphate content. Seeds, cereal grains and legumes are important feed ingredients.

Phytic acid, or its salts phytates - said terms being, unless otherwise indicated, in the present context used synonymously or at random - is an anti-nutritional factor. This is partly due to its binding of nutritionally essential ions such as calcium, trace minerals such as mangane, and also proteins (by electrostatic interaction). And partly due to the fact that the phosphorous thereof is not nutritionally available either, since phytic acid and its salts, phytates, are often not metabolized.

This leads to a need of supplementing food and feed preparations with e.g. inorganic phosphate.

The non-metabolizable phytic acid phosphorous passes through the gastrointestinal tract of such animals and is excreted with the manure, resulting in an undesirable phosphate pollution of the environment resulting e.g. in eutrophication of the water environment and extensive growth of algae.

Phytic acid is degradable by phytases. In the present context a "phytase" is an polypeptide or enzyme which exhibits phytase activity, viz. which catalyzes the hydrolysis of phytate (myo-inositol hexakisphosphate) to (1) myo-inositol and/or (2) mono-, di-, tri-, tetra- and/or penta-phosphates thereof and (3) inorganic phosphate.

The production of phytases by plants as well as by microorganisms has been reported. Amongst the microorganisms, phytase producing bacteria as well as phytase producing fungi are known.

There are several descriptions of phytase producing filamentous fungi belonging to the fungal phylum of Ascomycota (ascomycetes). In particular, there are several references to phytase producing ascomycetes of the Aspergillus genus such as Aspergillus terreus (Yamada et al., 1986, Agric. Biol. Chem. 322:1275-1282). Also, the cloning and expression of the phytase gene from Aspergillus niger var. awamori has been described (Piddington et al., 1993, Gene 133:55-62). EP 0420358 describes the cloning and expression of a phytase of Aspergillus ficuum (niger). EP 0684313 describes the cloning and expression of phytases of the ascomycetes Aspergillus niger, Myceliophthora thermophila, Aspergillus terreus. Still further, some partial sequences of phytases of Aspergillus nidulans, Talaromyces thermophilus, Aspergillus fumigatus and another strain of Aspergillus terreus are given.

The cloning and expression of a phytase of Thermomyces lanuginosus is described in WO 97/35017.

WO 98/28409 describes the cloning and expression of several basidiomycete phytases, e.g. from Peniophora lycii, Agrocybe pediades, Paxillus involutus and Trametes pubescens.

According to the Enzyme nomenclature database ExPASy (a repository of information relative to the nomenclature of enzymes primarily based on the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB) describing each type of characterized enzyme for which an EC (Enzyme Commission) number has been provided), two different types of phytases are presently known: A so-called 3-phytase (myo-inositol hexaphosphate 3-phosphohydrolase, EC 3.1.3.8) and a so-called 6-phytase (myo-inositol hexaphosphate 6-phosphohydrolase, EC 3.1.3.26). The 3-phytase hydrolyses first the ester bond at a 3-position, whereas the 6-phytase hydrolyzes first an ester bond at the 6-position of phytic acid. Both of these types of phytases are included in the above definition of phytase.

Many assays of phytase activity are known, and any of these can be used for the purpose of the present invention. Preferred phytase assays are included in the examples.

The concept of "agglomeration" is defined as a process in which various components are mixed under the influence of heat. The resulting product is preferably an "agglomerate" or conglomerate in which the components adhere to each other while forming a product of a satisfactory physical stability. The formation of dust from such agglomerate is an indication of its physical stability - the less dust being formed, the better. A suitable assay for dust formation from agglomerates is ASAE standard S 269-1. A satisfactory agglomerate has below 20%, preferably below 15%, more preferably below 10%, even more preferably below 6% dust.

"Under the influence of heat" means that the temperature is at least 65°C, as measured on the product at the outlet of the agglomeration unit. More preferred temperatures are at least 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, or even at least 130°C.

A preferred agglomeration process is operated at an increased pressure. The pressure is typically due to a compacting of the ingredients, optionally in combination with a reduction of the cross-sectional or throughput area. Preferably, by properly adjusting process parameters such as temperature and pressure, the resulting shear forces and shear velocities are of such magnitude, that the starch- and protein-containing feed ingredients become fluid.

"Increased pressure" means increased as compared to normal atmospheric pressure, and the maximum pressure as measured within the agglomeration unit.

The addition of water vapour or steam is often included in agglomeration, but not as an absolute requirement.

Agglomeration includes, but is not limited to, the well-known processes called extrusion, expansion (or pressure conditioning) and pelleting (or pellet pressing).

Extrusion is i.a. described at pp. 149-153 of a handbook which is available on request from the Danish Company Sprout-Matador, Glentevej 5-7, DK-6705 Esbjerg Ø or Niels Finsensvej 4, DK-7100 Vejle ("Håndbog i Pilleteringsteknik 1996"). However, in the agglomeration process of the invention, the following process steps mentioned in the above handbook are entirely optional:

(i) pre-treating the feed ingredients in a cascade mixer;
(ii) cutting the product leaving the nozzle-section into pieces
(iii) of a desired size;
(iv) acclimatizing or conditioning it;
(v) coating it;
(vi) drying it;
(vii) cooling it.

The process of expansion (pressure conditioning) is i.a. described in the same handbook at pp. 61-66. Also for expansion, the above process steps (i)-(vi), in particular steps (i) and (vi), are entirely optional steps.

This is so also for the following process steps:

(ii') comminuting the product (using e.g. a blade granulator as shown at p. 65);
(vii) pelleting the product (using e.g. a pellet press as shown at p. 62);

The process of pelleting is i.a. described in the same handbook at pp. 71-107. Also here, steps (i)-(vii) above are entirely optional steps. These steps are i.a. described in more detail at pp. 29-70 of the above handbook.

In a preferred agglomeration process of the invention, one or more of the above mentioned further process steps (i)-(vii) are included.

A particularly preferred further step is step (i).

In a most preferred embodiment, the feed-ingredients are pre-heated in a first step (a) to a temperature of at least 45°C, preferably at least 50, 55, 60, 65, 70, 75, 80 °C; and then heated in a second step (b) to a temperature of at least

65°C, preferably 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, or even at least 130°C.

The addition of thermostable phytase takes place before or during step (a) and/or before or during step (b).

Water is preferably added in step (a). More preferably, heated steam is added during the mixing of the ingredients (steps (a) and/or (b)).

Process step (a) is preferably performed in a cascade mixer (see the above cited handbook p. 44).

A "thermostable" phytase is a phytase which has a Tm (melting temperature) as measured on purified phytase protein by Differential Scanning Calorimetry (DSC) of at least 65°C, preferably using for the DSC a constant heating rate, more preferably of 10°C/min. In preferred embodiments, the Tm is at least 66, 67, 68, 69, 70, 71, 72, 73, 74 or 75°C. Preferably, the Tm is equal to or lower than 150°C, more preferably equal to or lower than 145, 140, 135, 130, 125, 120, 115 or 110°C. Accordingly, preferred intervals of Tm are: 65-150°C, 66-150°C, - (etc.) - 75-150°C; 65-145°C, 66-145°C, - (etc.) - 75-145°C; 65-140°C, - (etc.) - 75-140°C; - (etc.) - 65-110°C, 66-110°C, - (etc.) - 75-110°C.

Particularly preferred ranges for Tm are the following: between 65 and 110°C; between 70 and 110°C; between 70 and 100°C; between 75 and 95°C, or between 80 and 90°C.

In Example 3 below, the measurement of Tm by DSC is described, and the Tm's of a number of phytases are shown.

The optimum temperatures are also indicated, since - in the alternative - a thermostable phytase can be defined as a phytase having a temperature-optimum of at least 60°C. Preferably, the optimum temperature is determined on the substrate phytate at pH 5.5, or on the substrate phytic acid at pH 5.0. Preferred units are FYT, FTU or the units of Example 3. The phytase assay of Example 3 is most preferred.

In preferred embodiments, the optimum temperature is at least 61, 62, 63, 64, 65, 66, 67, 68, 69 or 70°C. Preferably, the optimum temperature is equal to or lower than 140°C, more preferably equal to or lower than 135, 130, 125, 120, 115, 110, 105 or 100°C. Accordingly, preferred intervals of optimum temperature are: 60-140°C, 61-140°C, - (etc.) - 70-140°C; 60-135°C, 61-135°C, - (etc.) - 70-135°C; 60-130°C, - (etc.) - 70-130°C; - (etc.) - 60-100°C, 61-100°C, - (etc.) - 70-100°C.

Preferred phytases of the present invention exhibit a degree of similarity or homology, preferably identity, to the complete amino acid sequence of either of the phytases mentioned below under (iii) - preferably to the complete amino acid sequence of Consensus-phytase-10-thermo-Q50T-K91A - of at least 48%, preferably at least 50, 52, 55, 60, 62, 65, 67, 70, 73, 75, 77, 80, 82, 85, 88, 90, 92, 95, 98 or 99%.

The degree of similarity or homology, alternatively identity, can be determined using any alignment programme known in the art. A preferred alignment programme is GAP provided in the GCG version 8 program package (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711) (see also Needleman, S.B. and Wunsch, C.D., (1970), Journal of Molecular Biology, 48, 443-453). Using GAP with the following settings for polypeptide sequence comparison: GAP weight of 3.000 and GAP lengthweight of 0.100.

A multiple sequence alignment can be made using the program PileUp (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science

Drive, Madison, Wisconsin, USA 53711), with a GapWeight of 3.000 and a GapLengthWeight of 0.100.

Using the program GAP, some selected phytases exhibit the following percentage similarity (identity in brackets) to the Consensus-phytase-10-thermo(3)-Q50T-K91A amino acid sequence:

| | |
|---|---|
| A. fumigatus ATCC-13073 α-mutant | 86.7% (81.8%) |
| Basidiomycet consensus | 64.1% (49.0%) |
| Consensus-phytase-1 | 98.7% (97.9%) |
| Consensus-phytase-10 | 96.6% (94.4%) |
| Consensus-phytase-1-thermo(8)-Q50T-K91A | 97.4% (95.5%) |
| Consensus-phytase-11 | 96.5% (94.2%) |
| Consensus-phytase-12 | 92.5% (89.9%) |
| Consensus-phytase-7 | 95.5% (93.4%) |

A "purified" phytase is essentially free of other non-phytase polypeptides, e.g. at least about 20% pure, preferably at least about 40% pure, more preferably about 60% pure, even more preferably about 80% pure, most preferably about 90% pure, and even most preferably about 95% pure, as determined by SDS-PAGE.

Preferred thermostable phytases are the so-called consensus phytases of EP 98113176.6 (EP 0897985), viz.

(i) any thermostable phytase which is obtainable by the processes described therein;

(ii) a phytase comprising the amino acid sequence shown in Fig. 2 thereof or any variant or mutein thereof, preferred muteins being those comprising the substitutions Q50L; Q50T; Q50G; Q50T-Y51N or Q50L-Y51N.

Other preferred thermostable phytases are

(iii) a thermostable phytase which comprises at least one of the following amino acid sequence (some of which are shown in Figs. 5-12 herein), preferably the following phytases: Consensus-phytase-1 (or simply Consensus phytase); Consensus-phytase-1-thermo(3); Consensus-phytase-1-Q50T; basidiomycete-consensus (or simply Basidio); Consensus-phytase-10 (or Fcp 10); Consensus-phytase-11 (or Consensus Seq. 11); Consensus-phytase-1-thermo(8)-Q50T-K91A; Consensus-phytase-1-thermo(8)-Q50T; Consensus-phytase-1-thermo(8); Consensus-phytase-10-thermo(3)-Q50T-K91A; Consensus-phytase-10-thermo(3)-Q50T (sometimes, "(3)" is deleted from this expression); Aspergillus fumigatus ATCC 13073 phytase α-mutant; Aspergillus fumigatus ATCC 13073 phytase α-mutant plus the mutations E59A, S126N, R329H, S364T, G404A; Aspergillus fumigatus ATCC 13073 phytase α-mutant plus the mutations E59A, K68A, S126N, R329H, S364T, G404A; Consensus-phytase-7; Consensus-phytase-12.

(iv) as well as thermostable variants and muteins of the phytases of (iv) and (v), in particular those comprising one or more of the following substitutions: Q50L,T,G; Q50L-Y51N; Q50T-Y51N.

The term "plant" is intended to include not only whole plants as such, but also plant parts or organs, such as leaves, seeds or grains, stem, root, tubers, flowers, callus, fruits etc.; tissues, cells, protoplats etc.; as well as any combinations or sub-combinations thereof. Plant tissue cultures and plant cell lines as well as plant protoplasts are specifically included herein.

The term "transgenic plant" is a plant as defined above, which has been genetically modified, as well as its progeny and propagating material thereof having retained the genetical modification. Preferably, the transgenic plant comprises at least one specific gene introduced into an ancestral plant by recombinant gene technology. The term is not confined to a single plant variety.

The invention relates to a transgenic plant which comprises a DNA-construct encoding a thermostable phytase.

In a preferred embodiment the transgenic plant is a plant grouping which is characterized in that it comprises a DNA-construct encoding a thermostable phytase. The members of this plant grouping may very well possess individuality, but are clearly distinguishable from other varieties by their common characteristic feature of the the thermostable phytase DNA-construct.

Accordingly, the present teaching is applicable to more than one plant variety. No naturally occuring plant varieties are included amongst the plants of the invention.

In another preferred embodiment the invention relates to a transgenic plant variety or a variant thereof; a transgenic plant species, a transgenic plant genus, a transgenic plant family, and/or a transgenic plant order. More preferably, plant varieties as such are disclaimed.

Any thermostable phytase may be used in the present invention, e.g. any wild-type phytases, genetically engineered phytases, consensus phytases, phytase muteins, and/or phytase variants. Genetically engineered phytases include, but are not limited to, phytases prepared by site-directed mutagenesis, gene shuffling, random mutagenesis, etc.

The nucleotide sequence encoding a wild-type thermostable phytase may be of any origin, including mammalian, plant and microbial origin and may be isolated from these sources by conventional methods. Preferably, the nucleotide sequence is derived from a microorganism, such as a fungus, e.g. a yeast or a filamentous fungus, or a bacterium. The DNA sequence encoding a thermostable phytase may be isolated from the cell producing it, using various methods well known in the art (see e.g. WO 98/28409 and EP 0897985).

The nucleotide sequence encoding a thermostable genetically engineered or consensus phytase, including muteins and variants thereof, may be prepared in any way, e.g. as described in Example 3 hereof and in EP 0897985.

In order to accomplish expression of the thermostable phytase in a plant of the invention the nucleotide sequence encoding the phytase is inserted into an expression construct containing regulatory elements or sequences capable of directing the expression of the nucleotide sequence and, if necessary or desired, to direct secretion of the gene product or targetting of the gene product to the seeds of the plant.

In order for transcription to occur the nucleotide sequence encoding the thermostable phytase is operably linked to a suitable promoter capable of mediating transcription in the plant in question. The promoter may be an inducible promoter or a constitutive promoter. Typically, an inducible promoter mediates transcription in a tissue-specific or growth-stage specific manner, whereas a constitutive promoter provides for sustained transcription in all cell tissues. An example of a suitable constitutive promoter useful for the present invention is the cauliflower mosaic virus 35 S promoter. Transcription initiation sequences from the tumor-inducing plasmid (Ti) of

Agrobacterium such as the octopine synthase, nopaline synthase, or mannopine synthase initiator, are further examples of preferred constitutive promoters.

Examples of suitable inducible promoters include a seed-specific promoter such as the promoter expressing alpha-amylase in wheat seeds (see Stefanov et al, Acta Biologica Hungarica. Vol. 42, No. 4 pp. 323-330 (1991), a promoter of the gene encoding a rice seed storage protein such as glutelin, prolamin, globulin or albumin (Wu et al., Plant and Cell Physiology Vol. 39, No. 8 pp. 885-889 (1998)), a Vicia faba promoter from the legumin B4 and the unknown seed protein gene from Vicia faba described by Conrad U. et al, Journal of Plant Physiology Vol. 152, No. 6 pp. 708-711 (1998), the storage protein napA promoter from Brassica napus, or any other seed specific promoter known in the art, eg as described in WO 91/14772.

In order to increase the expression of the thermostable phytase it is desirable that a promoter enhancer element is used. For instance, the promoter enhancer may be an intron which is placed between the promoter and the amylase gene. The intron may be one derived from a monocot or a dicot. For instance, the intron may be the first intron from the rice Waxy (Wx) gene (Li et al., Plant Science Vol. 108, No. 2, pp. 181-190 (1995)), the first intron from the maize Ubil (Ubiquitin) gene (Vain et al., Plant Cell Reports Vol. 15, No. 7 pp. 489-494 (1996)) or the first intron from the Actl (actin) gene. As an example of a dicot intron the chsA intron (Vain et al. op cit.) is mentioned. Also, a seed specific enhancer may be used for increasing the expression of the thermostable phytase in seeds. An example of a seed specific enhancer is the one derived from the beta-phaseolin gene encoding the major seed storage protein of bean

(Phaseolus vulgaris) disclosed by Vandergeest and Hall, Plant Molecular Biology Vol. 32, No. 4, pp. 579-588 (1996).

Also, the expression construct preferably contains a terminator sequence to signal transcription termination of the thermostable phytase gene such as the rbcS2′ and the nos3′ terminators.

To facilitate selection of successfully transformed plants, the expression construct should also preferably include one or more selectable markers, e.g. an antibiotic resistance selection marker or a selection marker providing resistance to a herbicide. One widely used selection marker is the neomycin phosphotransferase gene (NPTII) which provides kanamycin resistance. Examples of other suitable markers include a marker providing a measurable enzyme activity, e.g. dihydrofolate reductase, luciferase, and b-glucoronidase (GUS). Phosphinothricin acetyl transferase may be used as a selection marker in combination with the herbicide basta or bialaphos.

The transgenic plant of the invention may be prepared by methods known in the art. The transformation method used will depend on the plant species to be transformed and can be selected from any of the transformation methods known in the art such as Agrobacterium mediated transformation (Zambryski et al., EMBO Journal 2, pp 2143-2150, 1993), particle bombardment, electroporation (Fromm et al. 1986, Nature 319, pp 791-793), and virus mediated transformation. For transformation of monocots particle bombardment (ie biolistic transformation) of embryogenic cell lines or cultured embryos are preferred. Below, references are listed, which disclose various methods for transforming various plants: Rice (Cristou et al. 1991, Bio/Technology 9, pp. 957-962), Maize (Gordon-Kamm et al. 1990, Plant Cell 2, pp. 603-618), Oat (Somers et al. 1992, Bio/Technology 10, pp 1589-1594), Wheat (Vasil et al. 1991, Bio/Technology 10, pp. 667-674, Weeks et al. 1993, Plant Physiology 102, pp. 1077-1084) and Barley (Wan and Lemaux 1994, Plant Physiology 102, pp. 37-48, review Vasil 1994, Plant Mol. Biol. 25, pp 925-937).

More specifically, Agrobacterium mediated transformation is conveniently achieved as follows:

A vector system carrying the thermostable phytase is constructed. The vector system may comprise of one vector, but it can comprise of two vectors. In the case of two vectors the vector system is referred to as a binary vector system (Gynheung An et al.(1980), Binary Vectors, Plant Molecular Biology Manual A3, 1-19).

An Agrobacterium based plant transformation vector consists of replication origin(s) for both E.coli and Agrobacterium and a bacterial selection marker. A right and preferably also a left border from the Ti plasmid from Agrobacterium tumefaciens or from the Ri plasmid from Agrobacterium rhizogenes is nessesary for the transformation of the plant. Between the borders the expression construct is placed which contains the thermostable phytase gene and appropriate regulatory sequences such as promotor and terminator sequences. Additionally, a selection gene e.g. the neomycin phosphotransferase type II (NPTII) gene from transposon Tn5 and a reporter gene such as the GUS (betha-glucuronidase) gene is cloned between the borders. A disarmed Agrobacterium strain harboring a helper plasmid containing the virulens genes is transformed with the above vector. The transformed Agrobacterium strain is then used for plant transformation.

The invention also relates to a method of preparing a transgenic plant capable of expressing a thermostable phytase, said method comprising the steps of (i) isolating a nucleotide sequence encoding a thermostable phytase; (ii) inserting the nucleotide sequence of (i) in an expression construct capable of mediating the expression of the nucleotide sequence in a selected host plant; and (iii) transforming the selected host plant with the expression construct.

The above method in which "at least one" replaces "a," when used in relation to the thermostable phytase, is also within this invention.

This method is an essentially non-biological method.

Any plant may be a selected host plant. More specifically, the plant can be dicotyledonous or monocotyledonous, for short a dicot or a monocot. Of primary interest are such plants which are potential food or feed components. These plants may comprise phytic acid. Examples of monocot plants are grasses, such as meadow grass (blue grass, Poa), forage grass such as festuca, lolium, temperate grass, such as Agrostis, and cereals, e.g. wheat, oats, rye, barley, rice, sorghum and maize (corn).

Examples of dicot plants are legumes, such as lupins, pea, bean and soybean, and cruciferous (family Brassicaceae), such as cauliflower, oil seed rape and the closely related model organism Arabidopsis thaliana.

Of particular interest are monocotyledonous plants, in particular crops or cereal plants such as wheat (Triticum, e.g. aestivum), barley (Hardeum, e.g. vulgare), oats, rye, rice, sorghum and corn (Zea, e.g. mays).

Of further particular interest are dicotyledonous plants, such as those mentioned above.

In a preferred embodiment, the ancestral plant or host plant is per se a desired feed ingredient.

## Examples

### Example 1

### FYT-assay - for analyzing phytase enzyme preparations

The phytase activity can be measured using the following assay: 10 µl diluted enzyme samples (diluted in 0.1 M sodium acetate, 0.01 % Tween20, pH 5.5) are added into 250 µl 5 mM sodium phytate (Sigma) in 0.1 M sodium acetate, 0.01 % Tween20, pH 5.5 (pH adjusted after dissolving the sodium phytate; the substrate is preheated) and incubated for 30 minutes at 37°C. The reaction is stopped by adding 250 µl 10 % TCA and free phosphate is measured by adding 500 µl 7.3 g FeSO4 in 100 ml molybdate reagent (2.5 g $(NH_4)_6Mo_7O_{24}.4H_2O$ in 8 ml $H_2SO_4$ diluted to 250 ml). The absorbance at 750 nm is measured on 200 µl samples in 96 well microtiter plates. Substrate and enzyme blanks are included. A phosphate standard curve is also included (0-2 mM phosphate). 1 FYT equals the amount of enzyme that releases 1 µmol phosphate/min at the given conditions. This assay is preferred for phytase enzyme preparations (when not in admixture with other feed ingredients).

### Example 2

### FTU assay - for analyzing phytase in admixture with feed ingredients

One FTU is defined as the amount of enzym, which at standard conditions (37°C, pH 5,5; reaction time 60 minutes and start concentration of phytic acid 5 mM) releases phosphate equivalent to 1 µmol phosphate per minute.

1 FTU = 1 FYT

The FTU assay is preferred for phytase activity measurements on animal feed premixes and the like complex compositions.

Reagents /substrates

Extraction buffer for feed etc.

This buffer is also used for preparation of $PO_4$-standards and further dilution of premix samples.

0,22 M acetate buffer with Tween 20 pH 5,5

30 g sodium acetate trihydrate (MW = 136,08 g/mol) e.g. Merck Art 46267 per liter and 0,1 g Tween 20 e.g. Merck Art 22184 pr. liter are weighed out.

The sodium acetate is dissolved in demineralised water.

Tween 20 is added, and pH adjusted to 5,50 ± 0,05 with acetic acid.

Add demineralised water to total volume.

Extraction buffer for premix

0,22 M acetate buffer with Tween 20, EDTA, $PO_4^{3-}$ og BSA.

30 g sodium acetate trihydrate e.g. Merck Art 6267 per liter.

0,1 g Tween 20 e.g. Merck Art 22184 per liter.

30 g EDTA f.eks. Merck Art 8418 pr. liter.

20 g $Na_2HPO_4$, $2H_2O$ e.g. Merck Art 6580 per liter.

0,5 g BSA (Bovine Serum Albumine, e.g. Sigma Art A-9647 per liter.

The ingredients are dissolved in demineralised water, and pH is adjusted to 5,50 ± 0,05 with acetic acid.

Add demineralised water to total volume.

BSA is not stable, and must therefore be added the same day the buffer is used.

50 mM $PO_4^{3-}$stock solution

0,681 g KH2PO4 (MW = 136,09 g/mol) e.g. Merck Art 4873 is weighed out and dissolved in 100 ml 0,22 M sodium acetat with Tween, pH 5,5.

Storage stability: 1 week in refrigerator.

0,22 M acetate buffer pH 5,5 without Tween

This buffer is used for production of phytic acid substrate).

150 g sodium acetate trihydrate (MW = 136,08) e.g. Merck Art 6267 is weighed out and dissolved in demineralised water, and pH is adjusted with acetic acid to 5,50 ± 0,05.

Add demineralised water to 5000 ml.

Storage stability: 1 week at room temperature.

Phytic acid substrate; 5 mM phytic acid

The volume of phytic acid is calculated with allowance for the water content of the used batch.

If the water content is e.g. 8,4 % the following is obtained:

$$\frac{0{,}005\,mol/l \times 923{,}8\,g/mol}{(1 \div 0{,}084)} = 5{,}04\,g/l$$

Phytic acid (Na-salt) (MW = 923,8 g/mol) e.g. Sigma P-8810 is weighed out and dissolved in 0,22 M acetate buffer (without tween). Addition of (diluted) acetic acid increases the dissolution speed.

pH is adjusted to 5,50 ± 0,05 with acetic acid.

Add 0,22 M acetate buffer to total volume.

21,7 % nitric acid solution

For stop solution.

1 part concentrated (65%) nitric acid is mixed into 2 parts demineralised water.

Molybdate reagent

For stop solution.

100 g ammonium heptamolybdate tetrahydrate $(NH_4)6Mo_7O_{24},4H_2O$ e.g. Merck Art 1182 is dissolved in demineralised water. 10 ml 25 % $NH_3$ is added.

Add demineralised water to 1 liter.

0,24 % Ammonium vanadate

Bought from fra Bie & Berntsen.

Molybdat/vanadat stop solution

1 part vanadate solution (0,24 % ammonium vanadate) + 1 part molybdate solution are mixed. 2 parts 21,7 % nitric acid solution are added.

The solution is prepared not more than 2 hours before use, and the bottle is wrapped in tinfoil.

Samples

Frozen samples are defrosted in a refrigerator overnight.

Sample size for feed samples: At least 70 g, preferably 100 g.

Feed samples

Choose a solution volume which allows addition of buffer corresponding to 10 times the sample weight, e.g. 100 g is dissolved in 1000 ml 0,22 M acetate buffer with Tween, see enclosure 1. Round up to nearest solution volume.

If the sample size is approx. 100 g all the sample is ground in a coffee grinder and subsequently placed in tared beakers. The sample weight is noted. It is not necessary to grind not-pelleted samples. If a sample is too big to handle, it is sample split into parts of approx. 100 g.

Magnets are placed in the beakers and 0,22 M acetate buffer with Tween is added.

The samples are extracted for 90 minutes.

After extraction the samples rest for 30 minuts to allow for the feed to sediment. A 5 ml sample is withdrawn with a pipette. The sample is taken 2 - 5 cm under the surface of the solution and placed in a centrifuge glass, which is covered by a lid.

The samples are centrifuged for 10 minutes at 4000 rpm.

Premix samples

Choose a solution volume which allows addition of buffer corresponding to 10 times the sample weight. Round up to nearest solution volume.

If the samples have been weighed (50 - 100 g) all of the sample is placed in tared beakers. The sample weight is noted. If a sample is too big to handle, it is split into parts of approx. 100 g.

Magnets are placed in the beakers and 0,22 M acetate buffer with Tween, EDTA og $PO_4^{3-}$ is added.

The samples are extracted for 60 minutes.

After extraction the samples rest for 30 minutes to allow for the premix to sediment. A 5 ml sample is withdrawn with a pipette. The sample is taken 2 - 5 cm under the surface of the solution and placed in a centrifuge glass, which is covered by a lid.

The samples are centrifuged for 10 minutes at 4000 rpm.

Analysis

Extracts of feed samples are analysed directly.

Extracts of premix are diluted to approx. 1,5 FTU/g ($A_{415}$ (main sample) < 1,0 ).

0,22 M acetate buffer with Tween 20 is used for the dilution.

Main Samples

2 x 100 ml of the supernatant from the extracted and centrifuged samples are placed in marked glass test tubes and a magnet is placed in each tube.

When all samples are ready they are placed on a water bath with stirring. Temperature: 37 °C.

3,0 ml substrate is added.

Incubation for exactly 60 minutes after addition of substrate.

The samples are taken off the water bath and 2,0 ml stop solution is added (exactly 60 minutes after addition of substrate).

The samples are stirred for 1 minute or longer.

Feed samples are centrifuged for 10 minutes at 4000 rpm (It is not necessary to centrifuge premix samples).

Blind samples

100 ml of the supernatant from the extracted and centrifuged samples are placed in marked glass test tubes, and a magnet is placed in each tube.

2,0 ml stop solution is added to the samples.

3,0 ml substrate is added to the samples.

The samples are incubated for 60 minutes at room temperature.

The feed samples are centrifuged for 10 minutes at 4000 rpm (it is not necessary to centrifuge premix samples).

Standards

2 x 100 ml are taken from each of the 8 standards and also 4 x 100 ml 0,22 M acetate buffer (reagent blind).

$A_{415}$ is measured on all samples.

CALCULATION

FTU/g = $\mu$mol $PO_4^{3-}$ / (min * g (sample))

*C* g sample is weighed out (after grinding).

100 $\mu$l is taken from the extracted and centrifuged sample.

$PO_4^{3-}$ standard curve is linear.

From the regression curve for the $PO_4^{3-}$ standard the actual concentration of the sample is found (concentration in mM):

$[PO_4^{3-}] = (x - b) / a$ $x = A_{415}$ $a$ = *slope* $b$ = *intercept with y-axis*

$\mu$mol $PO_4^{3-}$/min = { $[PO_4^{3-}]$ (mM) × Vol (liter) × 1000 $\mu$mol/mmol } /t

*t* = incubation time in minutes.

*Vol* = sample volume in liter = 0,0001 liter

1000 = conversion factor from mmol to $\mu$mol

$$\text{FTU}/g_{prøve} = \{ (x - b) \times Vol \times 1000 \times F_p\} / \{ a \times t \times C\}$$

$C$ = gram sample weighed out

$F_P$ = Relation between the sample taken out and the total sample (after extraction). Example: 0,100 ml taken from 1000 ml → $F_P$ = 1000/0,100 = 10000.

Reduced expression with insertion of the following values:

$t$ = 60

$Vol$ = 0,0001 l

$F_P$ = 10000

$$\text{FTU}/g_{sample} = \{ (x - b) \times 0{,}0001 \times 1000 \times 10000\} / \{ a \times 60 \times C\}$$

## Example 3

## Determination of optimum temperature and melting point Tm of various phytases

The thermostability of various phytases has been determined, viz. the melting temperature, Tm, and/or the optimum temperature.

The phytase of Aspergillus niger NRRL 3135 was prepared as described in EP 0420358 and van Hartingsveldt et al (Gene, 127, 87-94, 1993).

The phytases of Aspergillus fumigatus ATCC 13073, Aspergillus terreus 9A-1, Aspergillus terreus CBS 116.46, Aspergillus nidulans, Myceliophthora thermophila, and Talaromyces thermophilus were prepared as described in EP-0897985 and the references therein.

Consensus-phytase-1 (Fig. 5) and Consensus-phytase-1-Q50T are shown in and were prepared as described in EP 0897985.

Consensus-phytase-10 was derived and prepared according to the teachings of EP-0897985 (Examples 1-2 and 3-7, respectively), however adding to the alignment at Fig. 1 thereof the phytase sequence of Thermomyces lanuginosa (Berka et al, Appl. Environ. Microbiol. 64, 4423-4427, 1998) and a basidiomycete consensus sequence (derivation described below), omitting the sequence of A.niger T213, and assigning a vote weight of 0.5 for the remaining A.niger phytase sequences. The derivation of the sequence of Consensus-phytase-10 is shown in Fig. 7.

The basidiomycete consensus sequence was also derived according to the principles of EP-0897985, viz. from the five basidiomycete phytases of WO 98/28409, starting with the first amino acid residue of the mature phytases (excluding signal peptide). A vote weight of 0.5 was assigned to the two Paxillus phytases, all other genes were used with a vote weight of 1.0 - see Fig. 6.

The muteins Consensus-phytase-10-thermo, Consensus-phytase-10-thermo-Q50T-K91A (Fig. 10) and Consensus-phytase-10-thermo-Q50T were prepared from consensus-phytase-10, in analogy to Examples 5-8 of EP-0897985, by introducing the three back-mutations K94A, V158I and A396S ("thermo(3)" or "thermo") and, where applicable, also the mutations Q50T or Q50T-K91A.

The muteins Consensus-phytase-1-thermo(8), Consensus-phytase-1-thermo(8)-Q50T-K91A (Fig. 9) and Consensus-phytase-1-thermo(8)-Q50T, were prepared from consensus-phytase-1, in analogy to Example 8 of EP-0897985, by introducing the eight mutations E58A, D197N, E267D, R291I, R329H, S364T, A379K and G404A ("thermo(8)") and, where applicable, also the mutations Q50T or Q50T-K91A.

Consensus-phytase-1-thermo(3) was prepared from consensus-phytase-1 by introduction of the three mutations K94A, V158I and A396S.

An Aspergillus fumigatus so-called α-mutant (with the mutations Q51(27)T, F55Y, V100I, F114Y, A243L, S265P, N294D) and the further muteins thereof shown in Table 1 were prepared as generally described above. The position numbering refers to Fig. 11 hereof, except for the number in parentheses which refers to the numbering used in EP 0897010.

DNA constructs encoding the above thermostable phytases can be prepared e.g. according to the teachings of EP 0897985. For expression thereof in plants, reference is made to the present description.

In order to determine the unfolding temperature or melting temperature, Tm, of a phytase, differential scanning calorimetry was applied as previously published by Brugger et al (1997): "Thermal denaturation of phytases and pH 2.5 acid phosphatase studied by differential scanning calorimetry," in The Biochemistry of phytate and phytase (eds. Rasmussen, S.K; Raboy, V.; Dalbøge, H. and Loewus, F.; Kluwer Academic Publishers).

Homogenous or purified phytase solutions of 50-60 mg/ml of protein are prepared, and extensively dialyzed against 10 mM sodium acetate, pH 5.0. A constant heating rate of 10°C/min is applied up to 90-95°C.

The results of Tm determinations on the above phytases are shown in Table 1 below; for selected phytases also in Figs. 1-4.

In Table 1 below, the optimum temperature of various phytases is also indicated. For this determination, phytase activity was determined basically as described by Mitchell et al (Microbiology 143, 245-252, 1997): The activity was measured in an assay mixture containing 0.5% phytic acid (~ 5 mM) in 200 mM sodium acetate, pH 5.0. After 15 min of incubation at 37°C, the reaction was stopped by addition of an equal volume of 15% trichloroacetic acid. The liberated phosphate was quantified by mixing 100 µl of the assay mixture with 900µl $H_2O$ and 1 ml of 0.6 M $H_2SO_4$, 2% ascorbic acid and 0.5% ammonium molybdate. Standard solutions of potassium phosphate were used as reference. One unit of enzyme activity was defined as the amount of enzyme that releases 1 µmol phosphate per minute at 37°C. The protein concentration was determined using the enzyme extinction coefficient at 280 nm calculated according to Pace et al (Prot.Sci. 4, 2411-2423, 1995): Consensus phytase, 1.101; consensus phytase 7, 1.068; consensus phytase 10, 1.039.

For determination of the temperature optimum, enzyme (100µl) and substrate solution (100µl) were pre-incubated for 5 min at the given temperature. The reaction was started by addition of the substrate solution to the enzyme. After 15 min incubation, the reaction was stopped with trichloroacetic acid and the amount of phosphate released was determined. Phytase-activity-versus-temperature is plotted, and the temperature optimum is determined as that temperature at which the acitivity reaches its maximum value.

Table 1

Temperature optimum and Tm for various phytases

| **Phytase** | **Optimum temperature (°C)** | **Tm (°C)** |
|---|---|---|
| Aspergillus niger NRRL 3135 | 55 | 63.3 |
| Aspergillus fumigatus ATCC 13073 | 55 | 62.5 |

| | | |
|---|---|---|
| Aspergillus terreus 9A-1 | 49 | 57.5 |
| Aspergillus terreus CBS 116.46 | 45 | 58.5 |
| Aspergillus nidulans | 45 | 55.7 |
| Myceliophthora thermophila | 55 | - |
| Talaromyces thermophilus | 45 | - |
| Consensus-phytase-10-thermo-Q50T-K91A | 82 | 89.3 |
| Consensus-phytase-10-thermo-Q50T | 82 | 88.6 |
| Consensus-phytase-10 | 80 | 85.4 |
| Consensus-phytase-1-thermo(8)-Q50T-K91A | - | 85.7 |
| Consensus-phytase-1-thermo(8)-Q50T | 78 | 84.7 |
| Consensus-phytase-1-thermo(8) | 81 | - |
| Consensus-phytase-1-thermo(8)-Q50T-K91A | 78 | 84.7 |
| Consensus-phytase-1-thermo(3) | 75 | - |
| Consensus-phytase-1-Q50T | - | 78.9 |
| Consensus-phytase-1 | 71 | 78.1 |
| Aspergillus fumigatus α-mutant, plus mutations E59A, | 63 | - |

| | | |
|---|---|---|
| S126N, R329H, S364T, G404A | | |
| Aspergillus fumigatus - as above, plus mutation K68A | 63 | – |
| Aspergillus fumigatus α-mutant (Q51(27)T, F55Y, V100I, F114Y, A243L, S265P, N294D) | 60 | 67.0 |

**The claims defining the invention are as follows:**

1. A process for preparing an animal feed, wherein a phytase with a Tm of at least 75°C is added before or during an agglomeration of feed ingredients, the Tm being measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE.

2. The process of claim 1, wherein the feed ingredients are heated to a temperature of at least 65°C.

3. The process of claim 1 or claim 2, when performed in a feed expander.

4. The process of claim 1 or claim 2, when performed in an extruder.

5. The process of claim 1 or claim 2, when performed in a pellet press.

6. The process of any one of claims 1 to 5, wherein the phytase is present in a transgenic plant.

7. The process of any one of claims 1 to 6, wherein the agglomeration includes the following steps:

(a) pre-heating the feed ingredients to a temperature of at least 45°C; and

(b) heating the product of step (a) to a temperature of at least 65°C; wherein the phytase is added prior to or during step (a) and/or (b).

8. The process of any one of claims 1 to 7, wherein the phytase is selected from the following group of phytases as described herein:

Consensus-phytase-1; Consensus-phytase-l-thermo (3); Consensus-phytase-1-Q50T; basidiomycete-consensus; Consensus phytase-10; Consensus-phytase-11; Consensus-phytase-l-thermo (8)-Q50T-K91A; Consensus-phytase-l-thermo (8)-Q50T; Consensus-phytase-1-thermo (8); Consensus-phytase-10-thermo (3)-Q50T-K91A; Consensus-phytase-10-thermo (3)-Q50T; Aspergillus fumigatus ATCC 13073 phytase α-mutant; Aspergillus fumigatus ATCC 13073 phytase α-mutant plus the mutations E59A, S126N, R329H, S364T, G404A; Aspergillus fumigatus ATCC 13073 phytase α-mutant plus the mutations E59A, K68A, S126N, R329H, S364T, G404A; Consensus-phytase-7; Consensus-phytase-12; and any thermostable variants and muteins thereof comprising one or more of the following substitutions: Q50L, T, G; Q50L-Y51N; or Q50T-Y51N.

9. The process of claim 8, wherein the phytase is selected from:
Consensus-phytase-1; Consensus-phytase-l-thermo (3); Consensus-phytase-1-Q50T; Consensus phytase-10; Consensus-phytase-l-thermo (8)-Q50T-K91A; Consensus-phytase-l-thermo (8)-Q50T; Consensus-phytase-1-thermo (8); Consensus-phytase-10-thermo (3)-Q50T-K91A; and Consensus-phytase-10-thermo (3)-Q50T.

10. The process of claim 9, wherein the phytase has the amino acid sequence of Consensus-phytase-l-thermo (8)-Q50T-K91A of Fig. 9 or the amino acid sequence of Consensus-phytase-l0-thermo (3)-Q50T-K91A of Fig. 10.

11. An animal feed obtainable by the process of any one of claims 1 to 10.

12. Use of a phytase with a Tm of at least 75°C in a process for preparing an animal feed, wherein the phytase is added before or during an agglomeration of feed ingredients, and wherein the Tm is measured by Differential Scanning Calorimetry (DSC)

using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE.

13. Use of a phytase with a Tm of at least 75°C in the manufacture of a preparation for use in a process for preparing an animal feed in which the phytase preparation is added before or during an agglomeration of feed ingredients, and wherein the Tm is measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE.

14. The use of claim 12 or claim 13, wherein in said process the feed ingredients are heated to a temperature of at least 65°C.

15. The use of claim 12 or claim 13, wherein said process is performed in a feed expander.

16. The use of claim 12 or claim 13, wherein said process is performed in an extruder.

17. The use of claim 12 or claim 13, wherein said process is performed in a pellet press.

18. The use of any one of claims 12 to 17, wherein the phytase is present in a transgenic plant.

19. The use of any one of claims 12 to 18, wherein the agglomeration includes the following steps:

(a) pre-heating the feed ingredients to a temperature of at least 45°C; and

(b) heating the product of step (a) to a temperature of at least 65°C; wherein the phytase is added prior to or during step (a) and/or (b).

20. The use of any one of claims 12 to 19, wherein the phytase is selected from the following group of phytases as described herein:

Consensus-phytase-1; Consensus-phytase-l-thermo (3); Consensus-phytase-1-Q50T; basidiomycete-consensus; Consensus phytase-10; Consensus-phytase-11; Consensus-phytase-l-thermo (8)-Q50T-K91A; Consensus-phytase-l-thermo (8)-Q50T; Consensus-phytase-1-thermo (8); Consensus-phytase-10-thermo (3)-Q50T-K91A; Consensus-phytase-10-thermo (3)-Q50T; Aspergillus fumigatus ATCC 13073 phytase α-mutant; Aspergillus fumigatus ATCC 13073 phytase α-mutant plus the mutations E59A, S126N, R329H, S364T, G404A; Aspergillus fumigatus ATCC 13073 phytase α-mutant plus the mutations E59A, K68A, S126N, R329H, S364T, G404A; Consensus-phytase-7; Consensus-phytase-12; and any thermostable variants and muteins thereof comprising one or more of the following substitutions: Q50L, T, G; Q50L-Y51N; or Q50T-Y51N.

21. The use of claim 20, wherein the phytase is selected from:
Consensus-phytase-1; Consensus-phytase-l-thermo (3); Consensus-phytase-1-Q50T; Consensus phytase-10; Consensus-phytase-l-thermo (8)-Q50T-K91A; Consensus-phytase-l-thermo (8)-Q50T; Consensus-phytase-1-thermo (8); Consensus-phytase-10-thermo (3)-Q50T-K91A; and Consensus-phytase-10-thermo (3)-Q50T.

22. The use of claim 21, wherein the phytase has the amino acid sequence of Consensus-phytase-l-thermo (8)-Q50T-K91A of Fig. 9 or the amino acid sequence of Consensus-phytase-l0-thermo (3)-Q50T-K91A of Fig. 10.

23. An expression construct which comprises a DNA construct encoding a phytase with a Tm of at least 75°C, the Tm being measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE, said DNA construct being operably linked to regulatory sequences capable of mediating expression of said phytase encoding sequence in at least one part of a plant.

24. The expression construct of claim 23, wherein the phytase is selected from the following group of phytases as described herein:

Consensus-phytase-1; Consensus-phytase-l-thermo (3); Consensus-phytase-1-Q50T; basidiomycete-consensus; Consensus phytase-10; Consensus-phytase-11; Consensus-phytase-l-thermo (8)-Q50T-K91A; Consensus-phytase-l-thermo (8)-Q50T; Consensus-phytase-1-thermo (8); Consensus-phytase-10-thermo (3)-Q50T-K91A; Consensus-phytase-10-thermo (3)-Q50T; Aspergillus fumigatus ATCC 13073 phytase α-mutant; Aspergillus fumigatus ATCC 13073 phytase α-mutant plus the mutations E59A, S126N, R329H, S364T, G404A; Aspergillus fumigatus ATCC 13073 phytase α-mutant plus the mutations E59A, K68A, S126N, R329H, S364T, G404A; Consensus-phytase-7; Consensus-phytase-12; and any thermostable variants and muteins thereof comprising one or more of the following substitutions: Q50L, T, G; Q50L-Y51N; or Q50T-Y51N.

25. The expression construct of claim 24, wherein the phytase is selected from:

Consensus-phytase-1; Consensus-phytase-l-thermo (3); Consensus-phytase-1-Q50T; Consensus phytase-10; Consensus-phytase-l-thermo (8)-Q50T-K91A; Consensus-phytase-l-thermo (8)-Q50T; Consensus-phytase-1-thermo (8); Consensus-phytase-10-thermo (3)-Q50T-K91A; and Consensus-phytase-10-thermo (3)-Q50T.

26. The expression construct of claim 25, wherein the DNA construct comprises a nucleotide sequence encoding the amino acid sequence of Fig. 9 or the nucleotide sequence of Fig. 10.

27. The expression construct of claim 26, wherein the DNA construct comprises the nucleotide sequence of Fig. 9 or a nucleotide sequence encoding Consensus-phytase-l0-thermo (3)-Q50T-K91A of Fig. 10.

28. A vector which comprises the expression construct of any one of claims 23 to 27.

29. The vector of claim 28, wherein the phytase is selected from the following group of phytases as described herein:

Consensus-phytase-1; Consensus-phytase-l-thermo (3); Consensus-phytase-1-Q50T; basidiomycete-consensus; Consensus phytase-10; Consensus-phytase-11; Consensus-phytase-l-thermo (8)-Q50T-K91A; Consensus-phytase-l-thermo (8)-Q50T; Consensus-phytase-1-thermo (8); Consensus-phytase-10-thermo (3)-Q50T-K91A; Consensus-phytase-10-thermo (3)-Q50T; Aspergillus fumigatus ATCC 13073 phytase α-mutant; Aspergillus fumigatus ATCC 13073 phytase α-mutant plus the mutations E59A, S126N, R329H, S364T, G404A; Aspergillus fumigatus ATCC 13073 phytase α-mutant plus the mutations E59A, K68A, S126N, R329H, S364T, G404A; Consensus-phytase-7; Consensus-phytase-12; and any thermostable variants and muteins thereof comprising one or more of the following substitutions: Q50L, T, G; Q50L-Y51N; or Q50T-Y51N.

30. The vector of claim 29, wherein the phytase is selected from:

Consensus-phytase-1; Consensus-phytase-l-thermo (3); Consensus-phytase-1-Q50T; Consensus phytase-10; Consensus-phytase-l-thermo (8)-Q50T-K91A; Consensus-phytase-l-thermo (8)-Q50T; Consensus-phytase-1-thermo (8); Consensus-phytase-10-thermo (3)-Q50T-K91A; and Consensus-phytase-10-thermo (3)-Q50T.

31. The vector of claim 30, wherein the DNA construct comprises a nucleotide sequence encoding the amino acid sequence of Fig. 9 or a nucleotide sequence encoding the amino acid sequence of Fig. 10.

32. The vector of claim 31, wherein the DNA construct comprises the nucleotide sequence of Fig. 9 or the nucleotide sequence of Fig. 10.

33. A method of preparing a transgenic plant capable of expressing a phytase with a Tm of at least 75°C, the Tm being measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE, said method comprising the steps of

(i) isolating a nucleotide sequence encoding the phytase;

(ii) inserting the nucleotide sequence of (i) in an expression construct capable of mediating the expression of the nucleotide sequence in a selected host plant; and

(iii) transforming the selected host plant with the expression construct.

34. The method of claim 33, wherein the expression construct is an expression construct of any one of claims 23 to 27.

35. The method of claim 34, wherein the expression construct is inserted into the host plant using a vector of any one of claims 28 to 32.

36. The method of any one of claims 33 to 35, which comprises the further step of extracting the phytase from the plant.

37. A method of preparing a transgenic plant cell capable of expressing a phytase with a Tm of at least 75°C, the Tm being measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE, said method comprising the steps of

(i) isolating a nucleotide sequence encoding the phytase;

(ii) inserting the nucleotide sequence of (i) in an expression construct capable of mediating the expression of the nucleotide sequence in a selected host plant cell; and

(iii) transforming the selected host plant cell with the expression construct.

38. The method of claim 37, wherein the expression construct is an expression construct of any one of claims 23 to 27.

39. The method of claim 38, wherein the expression construct is inserted into the host using a vector of any one of claims 28 to 32.

40. A transgenic plant cell capable of expressing a phytase with a Tm of at least 75°C, prepared by a method of any one of claims 37 to 39.

41. A transgenic plant capable of expressing a phytase with a Tm of at least 75°C, regenerated from a plant cell of claim 40.

42. A transgenic plant capable of expressing a phytase with a Tm of at least 75°C, comprising a plurality of cells of claim 40.

43. A transgenic plant which comprises a DNA-construct encoding a phytase with a Tm of at least 75°C, the Tm being measured by Differential Scanning Calorimetry (DSC) using a constant heating rate of 10°C/min, on a preparation of the phytase that has been dialysed against 10mM sodium acetate, pH 5.0, and is about 95% pure as determined by SDS-PAGE.

44. The transgenic plant of claim 43, wherein the DNA-construct encoding the phytase is operably linked to regulatory sequences capable of mediating expression of said phytase encoding sequence in at least one part of the plant.

45. The transgenic plant of claim 43 or claim 44, wherein said DNA construct is an expression construct of any one of claims 23 to 27.

46. A transgenic plant prepared by the method of any one of claims 33 to 35.

47. A method of producing a phytase with a Tm of at least 75°C, said method comprising culturing one or more transgenic cells of claim 40, or growing one or more transgenic plants of any one of claims 41 to 46, under conditions conducive to expression of said phytase in said plant cell, plant, or transgenic seed of said plant.

48. The method of claim 47, which comprises the further step of extracting the phytase from the plant cell, plant or transgenic seeds obtained therefrom.

49. A phytase with a Tm of at least 75°C produced by the method of claim 47 or claim 48.

50. The phytase of claim 49, when used in a process of any one of claims 1 to 10.

51. Transgenic seed from a plant of any one of claims 41 to 46.

**Dated 23 August, 2002**

**Novozymes A/S**

**Patent Attorneys for the Applicant/Nominated Person**

SPRUSON & FERGUSON

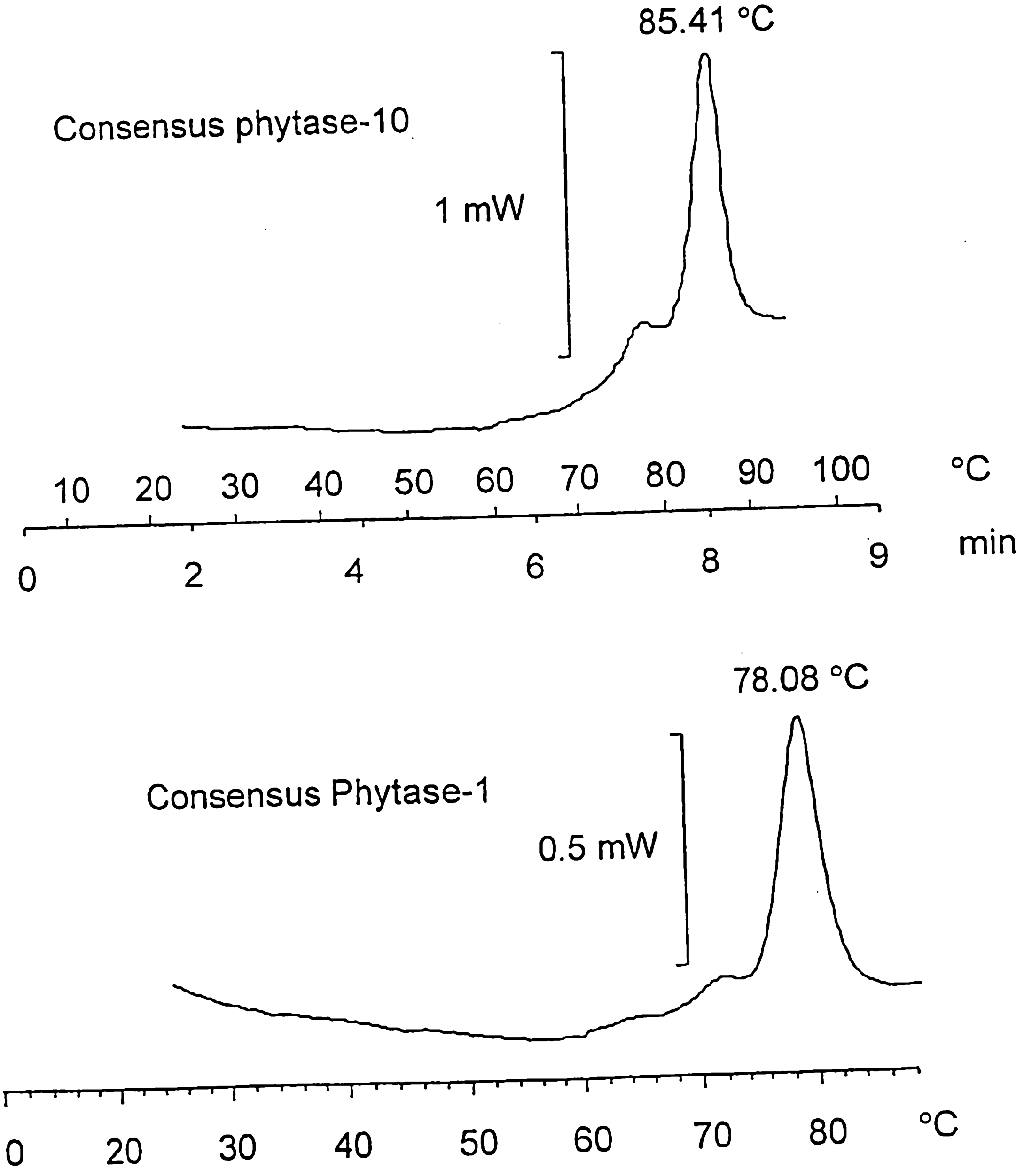
85.41 °C
Consensus phytase-10
1 mW
10 20 30 40 50 60 70 80 90 100 °C
0 2 4 6 8 9 min
78.08 °C
Consensus Phytase-1
0.5 mW
10 20 30 40 50 60 70 80 °C

Fig. 1

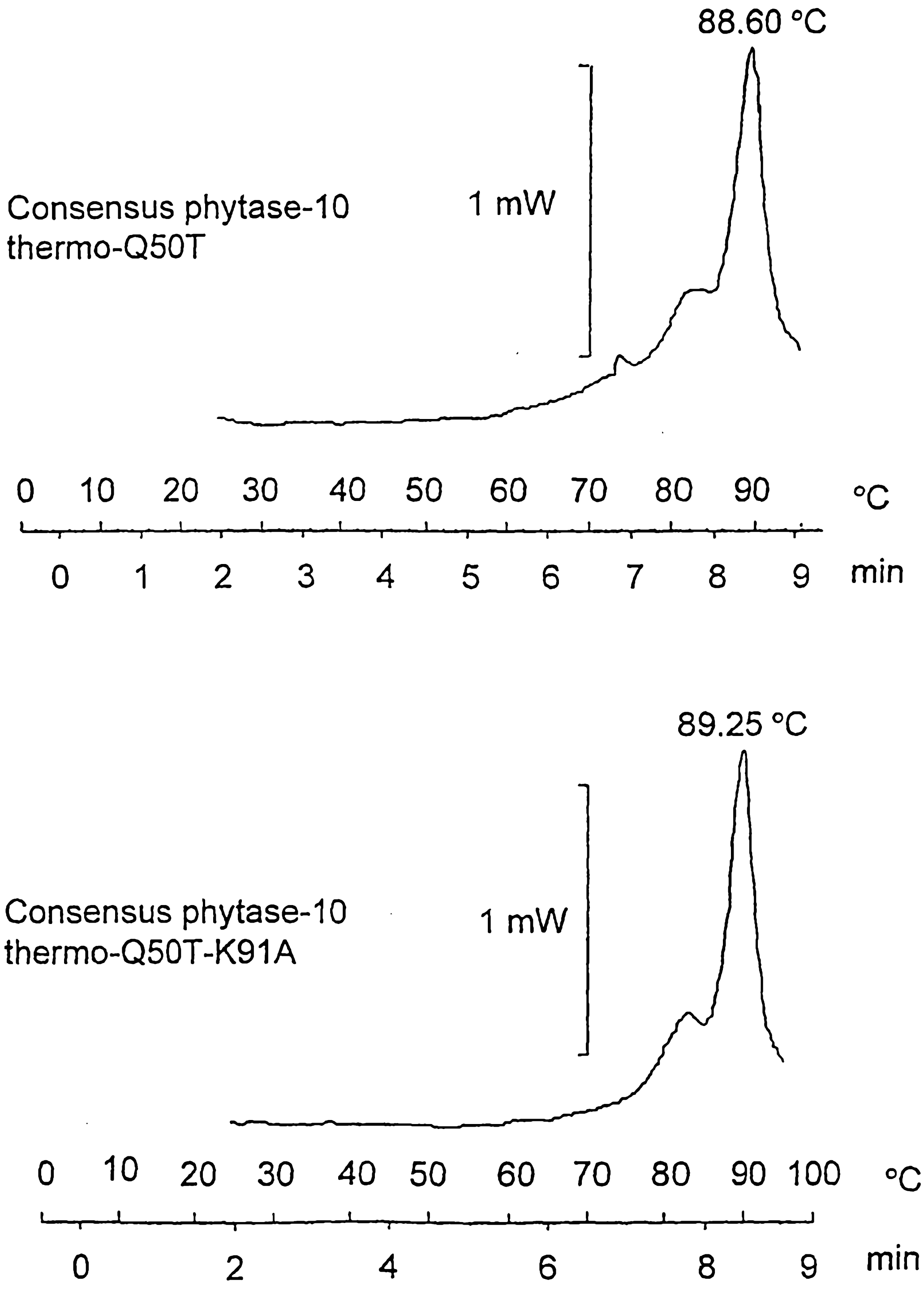
88.60 °C
Consensus phytase-10
thermo-Q50T
1 mW
0 10 20 30 40 50 60 70 80 90 °C
0 1 2 3 4 5 6 7 8 9 min
89.25 °C
Consensus phytase-10
thermo-Q50T-K91A
1 mW
0 10 20 30 40 50 60 70 80 90 100 °C
0 2 4 6 8 9 min

Fig. 2

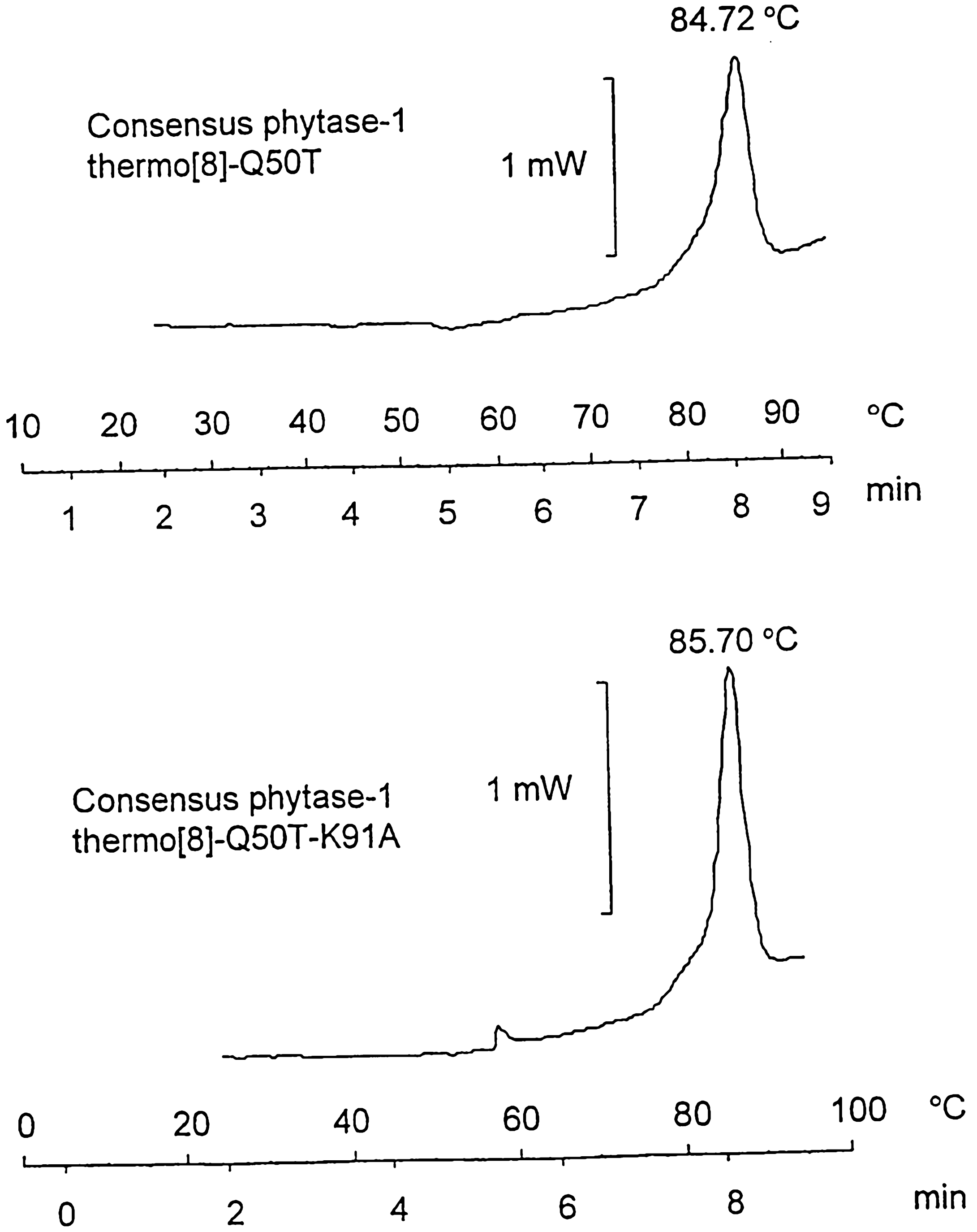
84.72 °C
Consensus phytase-1
thermo[8]-Q50T
1 mW
10 20 30 40 50 60 70 80 90 °C
1 2 3 4 5 6 7 8 9 min
85.70 °C
Consensus phytase-1
thermo[8]-Q50T-K91A
1 mW
0 20 40 60 80 100 °C
0 2 4 6 8 min

Fig. 3

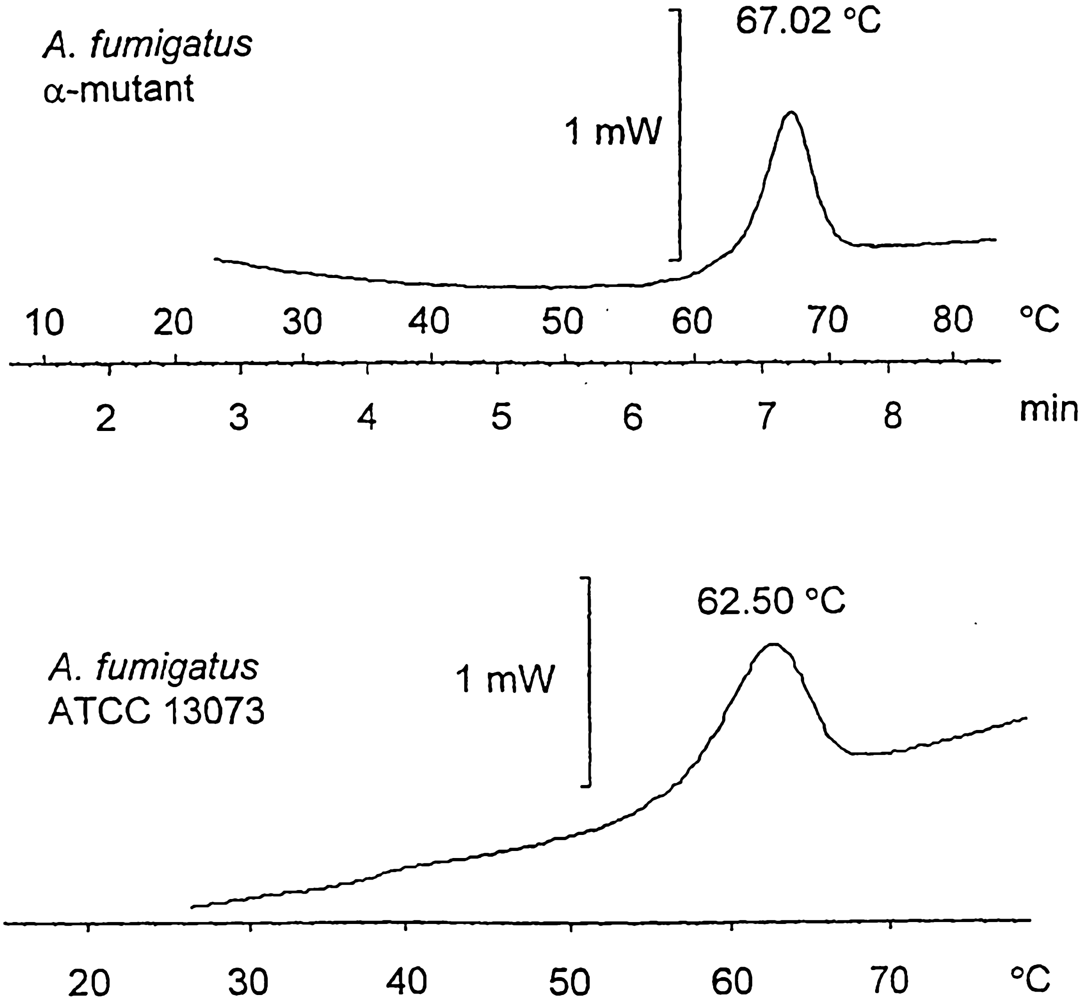
A. fumigatus
α-mutant
1 mW
67.02 °C
10
20
30
40
50
60
70
80
°C
2
3
4
5
6
7
8
min
A. fumigatus
ATCC 13073
1 mW
62.50 °C
20
30
40
50
60
70
°C

Fig. 4

```
                          1
50
A. terreus 9A-1           KhsDCNSVDh GYQCFPELSH kWGlYAPYFS LQDESPFPlD VPEDChITFV
A. terreus cbs            NhsDCTSVDr GYQCFPELSH kWGlYAPYFS LQDESPFPlD VPDDChITFV
A. niger var. awamori     NqsTCDTVDQ GYQCFSETSH LWGQYAPFFS LANESAISPD VPAGCrVTFA
A. niger T213             NqsSCDTVDQ GYQCFSETSH LWGQYAPFFS LANESVISPD VPAGCrVTFA
A. niger NRRL3135         NqsSCDTVDQ GYQCFSETSH LWGQYAPFFS LANESVISPE VPAGCrVTFA
A. fumigatus 13073        GSkSCDTVDl GYQCsPATSH LWGQYSPFFS LEDElSVSSK LPKDCrITLV
A. fumigatus 32722        GSkSCDTVDl GYQCsPATSH LWGQYSPFFS LEDElSVSSK LPKDCrITLV
A. fumigatus 58128        GSkSCDTVDl GYQCsPATSH LWGQYSPFFS LEDElSVSSK LPKDCrITLV
A. fumigatus 26906        GSkSCDTVDl GYQCsPATSH LWGQYSPFFS LEDElSVSSK LPKDCrITLV
A. fumigatus 32239        GSkACDTVEl GYQCsPGTSH LWGQYSPFFS LEDElSVSSD LPKDCrVTFV
E. nidulans               QNHSCNTADG GYQCFPNVSH VWGQYSPYFS IEQESAISeD VPHGCeVTFV
T. thermophilus           DSHSCNTVEG GYQCrPEISH sWGQYSPFFS LADQSEISPD VPQNCkITFV
M. thermophila            ESRPCDTpDl GFQCgTAISH FWGQYSPYFS VpSElDaS.. IPDDCeVTFA

Consensus                 NSHSCDTVDG GYQCFPEISH LWGQYSPYFS LEDESAISPD VPDDC-VTFV
Consensus phytase         NSHSCDTVDG GYQCFPEISH LWGQYSPYFS LEDESAISPD VPDDCRVTFV


                          51
100
A. terreus 9A-1           QVLARHGARs PThSKtKAYA AtIAAIQKSA TaFpGKYAFL QSYNYSLDSE
A. terreus cbs            QVLARHGARs PTDSKtKAYA AtIAAIQKNA TaLpGKYAFL KSYNYSMGSE
A. niger var. awamori     QVLSRHGARY PTESKgKkYS ALIEEIQQNV TtFDGKYAFL KTYNYSLGAD
A. niger T213             QVLSRHGARY PTESKgKkYS ALIEEIQQNV TtFDGKYAFL KTYNYSLGAD
A. niger NRRL3135         QVLSRHGARY PTDSKgKkYS ALIEEIQQNA TtFDGKYAFL KTYNYSLGAD
A. fumigatus 13073        QVLSRHGARY PTSSKsKkYK kLVTAIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus 32722        QVLSRHGARY PTSSKsKkYK kLVTAIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus 58128        QVLSRHGARY PTSSKsKkYK kLVTAIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus 26906        QVLSRHGARY PTSSKsKkYK kLVTAIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus 32239        QVLSRHGARY PTASKsKkYK kLVTAIQKNA TeFKGKFAFL ETYNYTLGAD
E. nidulans               QVLSRHGARY PTESKsKAYS GLIEAIQKNA TsFwGQYAFL ESYNYTLGAD
T. thermophilus           QLLSRHGARY PTSSKtElYS QLISrIQKTA TaYKGyYAFL KDYrYqLGAN
M. thermophila            QVLSRHGARa PTlKRaaSYv DLIDrIHhGA IsYgPgYEFL RTYDYTLGAD

Consensus                 QVLSRHGARY PTSSK-KAYS ALIEAIQKNA T-FKGKYAFL KTYNYTLGAD
Consensus phytase         QVLSRHGARY PTSSKSKAYS ALIEAIQKNA TAFKGKYAFL KTYNYTLGAD


                          101
150
A. terreus 9A-1           ELTPFGrNQL rDlGaQFYeR YNALTRhInP FVRATDASRV hESAEKFVEG
A. terreus cbs            NLTPFGrNQL qDlGaQFYRR YDTLTRhInP FVRAADSSRV hESAEKFVEG
A. niger var. awamori     DLTPFGEQEL VNSGIKFYQR YESLTRNIIP FIRSSGSSRV IASGEKFIEG
A. niger T213             DLTPFGEQEL VNSGIKFYQR YESLTRNIIP FIRSSGSSRV IASGEKFIEG
A. niger NRRL3135         DLTPFGEQEL VNSGIKFYQR YESLTRNIVP FIRSSGSSRV IASGKKFIEG
A. fumigatus 13073        DLTPFGEQQL VNSGIKFYQR YKALARSVVP FIRASGSDRV IASGEKFIEG
A. fumigatus 32722        DLTPFGEQQL VNSGIKFYQR YKALARSVVP FIRASGSDRV IASGEKFIEG
A. fumigatus 58128        DLTPFGEQQL VNSGIKFYQR YKALARSVVP FIRASGSDRV IASGEKFIEG
A. fumigatus 26906        DLTAFGEQQL VNSGIKFYQR YKALARSVVP FIRASGSDRV IASGEKFIEG
A. fumigatus 32239        DLTPFGEQQM VNSGIKFYQK YKALAgSVVP FIRSSGSDRV IASGEKFIEG
E. nidulans               DLTiFGENQM VDSGaKFYRR YKNLARKnTP FIRASGSDRV VASAEKFING
T. thermophilus           DLTPFGENQM IQlGIKFYnH YKSLARNaVP FVRCSGSDRV IASGrlFIEG
M. thermophila            ELTRtGQQQM VNSGIKFYRR YRALARKsIP FVRTAGqDRV VhSAENFTQG

Consensus                 DLTPFGENQM VNSGIKFYRR YKALARK-VP FVRASGSDRV IASAEKFIEG
Consensus phytase         DLTPFGENQM VNSGIKFYRR YKALARKIVP FIRASGSDRV IASAEKFIEG
```

Fig. 5A

```
                          151
200
A. terreus 9A-1           FQTARqDDHh ANpHQPSPrV DVaIPEGSAY NNTLEHSlCT AFES...STV
A. terreus cbs            FQNARqGDPh ANpHQPSPrV DVVIPEGTAY NNTLEHSICT AFEA...STV
A. niger var. awamori     FQSTKLkDPr AqpgQSSPkI DVVISEASSs NNTLDPGTCT VFED...SEL
A. niger T213             FQSTKLkDPr AqpgQSSPkI DVVISEASSs NNTLDPGTCT VFED...SEL
A. niger NRRL3135         FQSTKLkDPr AqpgQSSPkI DVVISEASSs NNTLDPGTCT VFED...SEL
A. fumigatus 13073        FQqAKLADPG A.TNRAAPAI SVIIPESETF NNTLDHGVCT kFEA...SQL
A. fumigatus 32722        FQqAKLADPG A.TNRAAPAI SVIIPESETF NNTLDHGVCT kFEA...SQL
A. fumigatus 58128        FQqAKLADPG A.TNRAAPAI SVIIPESETF NNTLDHGVCT kFEA...SQL
A. fumigatus 26906        FQqAKLADPG A.TNRAAPAI SVIIPESETF NNTLDHGVCT kFEA...SQL
A. fumigatus 32239        FQqANVADPG A.TNRAAPVI SVIIPESETY NNTLDHSVCT NFEA...SEL
E. nidulans               FRKAQLhDHG S..gQATPVV NVIIPEiDGF NNTLDHSTCV SFEN...DEr
T. thermophilus           FQSAKVlDPh SDkHDAPPTI NVIIeEGPSY NNTLDtGSCP VFED...SSg
M. thermophila            FHSAlLADRG STvRPTlPyd mVVIPETAGa NNTLHNDlCT AFEEgpySTI

Consensus                 FQSAKLADPG S-PHQASPVI NVIIPEGSGY NNTLDHGTCT AFED---SEL
Consensus phytase         FQSAKLADPG SQPHQASPVI DVIIPEGSGY NNTLDHGTCT AFED...SEL


                          201
250
A. terreus 9A-1           GDDAvANFTA VFAPAIaQRL EADLPGVqLS TDDVVnLMAM CPFETVSlTD
A. terreus cbs            GDAAADNFTA VFAPAIakRL EADLPGVqLS ADDVVnLMAM CPFETVSlTD
A. niger var. awamori     ADTVEANFTA TFAPSIRQRL ENDLSGVTLT DTEVTyLMDM CSFDTIStST
A. niger T213             ADTVEANFTA TFAPSIRQRL ENDLSGVTLT DTEVTyLMDM CSFDTIStST
A. niger NRRL3135         ADTVEANFTA TFVPSIRQRL ENDLSGVTLT DTEVTyLMDM CSFDTIStST
A. fumigatus 13073        GDEVAANFTA lFAPDIRARa EkHLPGVTLT DEDVVsLMDM CSFDTVARTS
A. fumigatus 32722        GDEVAANFTA lFAPDIRARa EkHLPGVTLT DEDVVsLMDM CSFDTVARTS
A. fumigatus 58128        GDEVAANFTA lFAPDIRARa EkHLPGVTLT DEDVVsLMDM CSFDTVARTS
A. fumigatus 26906        GDEVAANFTA lFAPDIRARa KkHLPGVTLT DEDVVsLMDM CSFDTVARTS
A. fumigatus 32239        GDEVEANFTA lFAPAIRARI EkHLPGVqLT DDDVVsLMDM CSFDTVARTA
E. nidulans               ADEiEANFTA IMGPPIRkRL ENDLPGIKLT NENVIyLMDM CSFDTMARTA
T. thermophilus           GHDAQEKFAk qFAPAIlEKI KDHLPGVDLA vSDVpyLMDL CPFETLARNh
M. thermophila            GDDAQDTYlS TFAGPItARV NANLPGANLT DADTVaLMDL CPFETVAsSS

Consensus                 GDDAEANFTA TFAPAIRARL EADLPGVTLT DEDVV-LMDM CPFETVARTS
Consensus phytase         GDDVEANFTA LFAPAIRARL EADLPGVTLT DEDVVYLMDM CPFETVARTS


                          251
300
A. terreus 9A-1           ......... ...DAhTLSPFC DLFTAtEWtq YNYLlSLDKY YGYGGGNPLG
A. terreus cbs            ......... ...DAhTLSPFC DLFTAaEWtq YNYLlSLDKY YGYGGGNPLG
A. niger var. awamori     ......... ...vDTKLSPFC DLFTHdEWih YDYLQSLkKY YGHGAGNPLG
A. niger T213             ......... ...vDTKLSPFC DLFTHdEWih YDYLRSLkKY YGHGAGNPLG
A. niger NRRL3135         ......... ...vDTKLSPFC DLFTHdEWin YDYLQSLkKY YGHGAGNPLG
A. fumigatus 13073        ......... ...DASQLSPFC QLFTHnEWkk YNYLQSLGKY YGYGAGNPLG
A. fumigatus 32722        ......... ...DASQLSPFC QLFTHnEWkk YNYLQSLGKY YGYGAGNPLG
A. fumigatus 58128        ......... ...DASQLSPFC QLFTHnEWkk YNYLQSLGKY YGYGAGNPLG
A. fumigatus 26906        ......... ...DASQLSPFC QLFTHnEWkk YNYLQSLGKY YGYGAGNPLG
A. fumigatus 32239        ......... ...DASELSPFC AIFTHnEWkk YDYLQSLGKY YGYGAGNPLG
E. nidulans               ......... ...HGTELSPFC AIFTEkEWlq YDYLQSLSKY YGYGAGSPLG
T. thermophilus           ......... ...TDT.LSPFC ALsTQeEWqa YDYYQSLGKY YGnGGGNPLG
M. thermophila            sdpatadagg gNGrpLSPFC rLFSEsEWra YDYLQSVGKW YGYGPGNPLG

Consensus                 ---------- -DATELSPFC ALFTE-EW-- YDYLQSLGKY YGYGAGNPLG
Consensus phytase         .......... .DATELSPFC ALFTHDEWRQ YDYLQSLGKY YGYGAGNPLG
```

Fig. 5B

```
                           301
350
A. terreus 9A-1            PVQGVGWaNE LMARLTRAPV HDHTCVNNTL DASPATFPLN ATLYADFSHD
A. terreus cbs             PVQGVGWaNE LIARLTRSPV HDHTCVNNTL DANPATFPLN ATLYADFSHD
A. niger var. awamori      PTQGVGYaNE LIARLTHSPV HDDTSSNHTL DSNPATFPLN STLYADFSHD
A. niger T213              PTQGVGYaNE LIARLTHSPV HDDTSSNHTL DSNPATFPLN STLYADFSHD
A. niger NRRL3135          PTQGVGYaNE LIARLTHSPV HDDTSSNHTL DSSPATFPLN STLYADFSHD
A. fumigatus 13073         PAQGIGFtNE LIARLTRSPV QDHTSTNsTL vSNPATFPLN ATMYVDFSHD
A. fumigatus 32722         PAQGIGFtNE LIARLTRSPV QDHTSTNsTL vSNPATFPLN ATMYVDFSHD
A. fumigatus 58128         PAQGIGFtNE LIARLTRSPV QDHTSTNsTL vSNPATFPLN ATMYVDFSHD
A. fumigatus 26906         PAQGIGFtNE LIARLTRSPV QDHTSTNsTL vSNPATFPLN ATMYVDFSHD
A. fumigatus 32239         PAQGIGFtNE LIARLTNSPV QDHTSTNsTL DSDPATFPLN ATIYVDFSHD
E. nidulans                PAQGIGFtNE LIARLTQSPV QDNTSTNHTL DSNPATFPLD rKLYADFSHD
T. thermophilus            PAQGVGFvNE LIARMTHSPV QDYTTVNHTL DSNPATFPLN ATLYADFSHD
M. thermophila             PTQGVGFvNE LLARLAgvPV RDgTSTNRTL DGDPrTFPLG rPLYADFSHD

Consensus                  PAQGVGF-NE LIARLTHSPV QDHTSTNHTL DSNPATFPLN ATLYADFSHD
Consensus phytase          PAQGVGFANE LIARLTRSPV QDHTSTNHTL DSNPATFPLN ATLYADFSHD


                           351
400
A. terreus 9A-1            SNLVSIFWAL GLYNGTAPLS qTSVESVSQT DGYAAAWTVP FAARAYVEMM
A. terreus cbs             SNLVSIFWAL GLYNGTkPLS qTTVEDITrT DGYAAAWTVP FAARAYIEMM
A. niger var. awamori      NGIISILFAL GLYNGTkPLS TTTVENITQT DGFSSAWTVP FASRlYVEMM
A. niger T213              NGIISILFAL GLYNGTkPLS TTTVENITQT DGFSSAWTVP FASRlYVEMM
A. niger NRRL3135          NGIISILFAL GLYNGTkPLS TTTVENITQT DGFSSAWTVP FASRlYVEMM
A. fumigatus 13073         NSMVSIFFAL GLYNGTEPLS rTSVESaKEl DGYSASWVVP FGARAYFEtM
A. fumigatus 32722         NSMVSIFFAL GLYNGTGPLS rTSVESaKEl DGYSASWVVP FGARAYFEtM
A. fumigatus 58128         NSMVSIFFAL GLYNGTEPLS rTSVESaKEl DGYSASWVVP FGARAYFEtM
A. fumigatus 26906         NSMVSIFFAL GLYNGTEPLS rTSVESaKEl DGYSASWVVP FGARAYFEtM
A. fumigatus 32239         NGMIPIFFAM GLYNGTEPLS qTSeESTKES NGYSASWAVP FGARAYFEtM
E. nidulans                NSMISIFFAM GLYNGTQPLS mDSVESIQEm DGYAASWTVP FGARAYFELM
T. thermophilus            NTMTSIFaAL GLYNGTAkLS TTEIKSIEET DGYSAAWTVP FGGRAYIEMM
M. thermophila             NDMMGVLgAL GaYDGVPPLD KTArrDpEEl GGYAASWAVP FAARiYVEKM

Consensus                  NSMISIFFAL GLYNGTAPLS TTSVESIEET DGYAASWTVP FGARAYVEMM
Consensus phytase          NSMISIFFAL GLYNGTAPLS TTSVESIEET DGYSASWTVP FGARAYVEMM


                           401
450
A. terreus 9A-1            QC........ .....RAEKE PLVRVLVNDR VMPLHGCPTD KLGRCKrDAF
A. terreus cbs             QC........ .....RAEKQ PLVRVLVNDR VMPLHGCAVD NLGRCKrDDF
A. niger var. awamori      QC........ .....QAEQE PLVRVLVNDR VVPLHGCPID aLGRCTrDSF
A. niger T213              QC........ .....QAEQE PLVRVLVNDR VVPLHGCPID aLGRCTrDSF
A. niger NRRL3135          QC........ .....QAEQE PLVRVLVNDR VVPLHGCPVD aLGRCTrDSF
A. fumigatus 13073         QC........ .....KSEKE PLVRALINDR VVPLHGCDVD KLGRCKLNDF
A. fumigatus 32722         QC........ .....KSEKE PLVRALINDR VVPLHGCDVD KLGRCKLNDF
A. fumigatus 58128         QC........ .....KSEKE SLVRALINDR VVPLHGCDVD KLGRCKLNDF
A. fumigatus 26906         QC........ .....KSEKE PLVRALINDR VVPLHGCDVD KLGRCKLNDF
A. fumigatus 32239         QC........ .....KSEKE PLVRALINDR VVPLHGCAVD KLGRCKLKDF
E. nidulans                QC........ .....E.KKE PLVRVLVNDR VVPLHGCAVD KFGRCTLDDW
T. thermophilus            QC........ .....DDSDE PVVRVLVNDR VVPLHGCEVD SLGRCKrDDF
M. thermophila             RCsggggggg ggegrQEKDE eMVRVLVNDR VMTLkGCGAD ErGMCTLErF

Consensus                  QC-------- -----QAEKE PLVRVLVNDR VVPLHGCAVD KLGRCKLDDF
Consensus phytase          QC........ .....QAEKE PLVRVLVNDR VVPLHGCAVD KLGRCKRDDF

                                  451
471
A. terreus 9A-1            VAGLSFAQAG GNWADCF--- -
A. terreus cbs             VEGLSFARAG GNWAECF--- -
A. niger var. awamori      VrGLSFARSG GDWAECsA-- -
A. niger T213              VrGLSFARSG GDWAECFA-- -
A. niger NRRL3135          VrGLSFARSG GDWAECFA-- -
A. fumigatus 13073         VKGLSWARSG GNWGECFS-- -
A. fumigatus 32722         VKGLSWARSG GNWGECFS-- -
A. fumigatus 58128         VKGLSWARSG GNWGECFS-- -
A. fumigatus 26906         VKGLSWARSG GNWGECFS-- -
A. fumigatus 32239         VKGLSWARSG GNSEQSFS-- -
E. nidulans                VEGLNFARSG GNWkTCFTl- -
```

Fig. 5C

```
T. thermophilus       VrGLSFARqG GNWEGCYAas e
M. thermophila        IESMAFARGN GKWDlCFA-- -

Consensus             VEGLSFARSG GNWAECFA-- -
Consensus phytase     VEGLSFARSG GNWAECFA.. .
```

Fig. 5D

```
                     1                                                     50
P. involutus (phyA1) SvP.KnTAPt FPIPeseQrn WSPYSPYFPL AeYkAPPAGC QInQVNIIQR
P. involutus (phyA2) SvP.RniAPK FSIPeseQrn WSPYSPYFPL AeYkAPPAGC EInQVNIIQR
T. pubescens         hiPlRdTSAc LdVTrDvQqs WSmYSPYFPa AtYvAPPASC QInQVHIIQR
A. pediades          GgvvQaTfvQ pfFPpQiQds WAAYTPYYPV qaYtPPPkDC KItQVNIIQR
P. lycii             StQfsfvAAQ LPIPaQntsn WGPYdPFFPV EpYaAPPEGC tVtQVNLIQR

Basidio              S-P-R-TAAQ LPIP-Q-Q-- WSPYSPYFPV A-Y-APPAGC QI-QVNIIQR


                     51                                                   100
P. involutus (phyA1) HGARFPTSGA TTRIKAGLTK LQGvqnfTDA KFNFIkSfkY dLGnsDLVPF
P. involutus (phyA2) HGARFPTSGA ATRIKAGLSK LQSvqnfTDP KFDFIkSfTY dLGtsDLVPF
T. pubescens         HGARFPTSGA AkRIQTAVAK LKAAsnyTDP lLAFVtNyTY sLGqDsLVeL
A. pediades          HGARFPTSGA GTRIQAAVkK LQSAktyTDP RLDFLtNyTY tLGhDDLVPF
P. lycii             HGARWPTSGA rSRqvAAVAK IQmArpfTDP KYEFLnDfvY kFGvADLLPF

Basidio              HGARFPTSGA ATRIQAAVAK LQSA---TDP KLDFL-N-TY -LG-DDLVPF


                     101                                                  150
P. involutus (phyA1) GAaQSfDAGQ EAFARYSkLV SkNNLPFIRA dGSDRVVDSA TNWTAGFAsA
P. involutus (phyA2) GAaQSfDAGl EvFARYSkLV SsDNLPFIRS dGSDRVVDTA TNWTAGFAsA
T. pubescens         GAtQSSEAGQ EAFTRYSsLV SaDELPFVRA SGSDRVVATA nNWTAGFAlA
A. pediades          GAlQSSQAGE ETFqRYSfLV SkENLPFVRA SSSNRVVDSA TNWTEGFSaA
P. lycii             GAnQShQTGt DmYTRYStLf egGDVPFVRA AGdQRVVDSS TNWTAGFGdA

Basidio              GA-QSSQAGQ EAFTRYS-LV S-DNLPFVRA SGSDRVVDSA TNWTAGFA-A


                     151                                                  200
P. involutus (phyA1) ShNTvqPkLn LILPQtGNDT LEDNMCPaAG DSDPQvNaWL AVafPSITAR
P. involutus (phyA2) SrNAiqPkLd LILPQtGNDT LEDNMCPaAG ESDPQvDaWL AsafPSVTAQ
T. pubescens         SsNSitPvLs VIISEaGNDT LDDNMCPaAG DSDPQvNqWL AqFAPPMTAR
A. pediades          ShHvlnPiLf VILSEslNDT LDDaMCPnAG sSDPQtGiWt SIYGTPIAnR
P. lycii             SgETvlPtLq VVLqEeGNcT LcNNMCPnEv DGDest.tWL GVFAPnITAR

Basidio              S-NT--P-L- VILSE-GNDT LDDNMCP-AG DSDPQ-N-WL AVFAPPITAR


                     201                                                  250
P. involutus (phyA1) LNAAAPSvNL TDtDAfNLvs LCAFlTVSkE kkSdFCtLFE giPGsFeAFa
P. involutus (phyA2) LNAAAPGANL TDaDAfNLvs LCPFmTVSkE qkSdFCtLFE giPGsFeAFa
T. pubescens         LNAGAPGANL TDtDTyNLlt LCPFETVAtE rrSeFCDIYE elQAE.dAFa
A. pediades          LNqqAPGANI TAaDvsNLip LCAFETIvkE tpSpFCNLF. .tPEEFaqFe
P. lycii             LNAAAPSANL SDsDAltLmd MCPFDTLSsG naSpFCDLF. .tAEEYvSYe

Basidio              LNAAAPGANL TD-DA-NL-- LCPFETVS-E --S-FCDLFE --PEEF-AF-


                     251                                                  300
P. involutus (phyA1) YgGDLDKFYG TGYGQeLGPV QGVGYVNELI ARLTnsAVRD NTQTNRTLDA
P. involutus (phyA2) YaGDLDKFYG TGYGQALGPV QGVGYINELL ARLTnsAVnD NTQTNRTLDA
T. pubescens         YnADLDKFYG TGYGQPLGPV QGVGYINELI ARLTaQnVsD HTQTNsTLDS
A. pediades          YfGDLDKFYG TGYGQPLGPV QGVGYINELL ARLTemPVRD NTQTNRTLDS
P. lycii             YyyDLDKYYG TGpGNALGPV QGVGYVNELL ARLTgQAVRD ETQTNRTLDS

Basidio              Y-GDLDKFYG TGYGQPLGPV QGVGYINELL ARLT-QAVRD NTQTNRTLDS


                     301                                                  350
```

Fig. 6A

```
P. involutus (phyA1) SPvTFPLNKT FYADFSHDNl MVAVFSAMGL FrQPAPLsTS vPNPwRTWrT
P. involutus (phyA2) APdTFPLNKT MYADFSHDNl MVAVFSAMGL FrQSAPLsTS tPDPNRTWLT
T. pubescens         SPeTFPLNRT LYADFSHDNQ MVAIFSAMGL FNQSAPLDPT tPDPaRTFLv
A. pediades          SPlTFPLDRS IYADLSHDNQ MIAIFSAMGL FNQSSPLDPS fPNPKRTWVT
P. lycii             dPaTFPLNRT FYADFSHDNt MVPIFAALGL FNaTA.LDPl kPDeNRlWVd

Basidio              SP-TFPLNRT FYADFSHDNQ MVAIFSAMGL FNQSAPLDPS -PDPNRTWVT


                     351                                                400
P. involutus (phyA1) SsLVPFSGRM VVERLsC..f GT........ .......tkV RVLVQDqVQP
P. involutus (phyA2) SsVVPFSARM aVERLsC..a GT........ .......tkV RVLVQDqVQP
T. pubescens         kKIVPFSARM VVERLdC..g GA........ .......qsV RLLVNDAVQP
A. pediades          SRLtPFSARM VtERLlCqrd GTgsggpsri mrngnvqtfV RILVNDALQP
P. lycii             SKLVPFSGHM tVEKLaC... .......... ....sgkeaV RVLVNDAVQP

Basidio              SKLVPFSARM VVERL-C--- GT-------- ---------V RVLVNDAVQP


                     401                                     441
P. involutus (phyA1) LEFCGGDrNG lCTLAkFVES QtFARsDGaG DFEKCFATSa -
P. involutus (phyA2) LEFCGGDqDG lCALDkFVES QaYARsGGaG DFEKCLATTv -
T. pubescens         LAFCGADtsG vCTLDAFVES QaYARNDGEG DFEKCFAT-- -
A. pediades          LKFCGGDmDS lCTLEAFVES QkYAREDGQG DFEKCFD--- -
P. lycii             LEFCGG.vDG vCeLsAFVES QtYARENGQG DFAKCgfvPs e

Basidio              LEFCGGD-DG -CTLDAFVES Q-YAREDGQG DFEKCFATP- -
```

Fig. 6B

```
                      1                                                   50
A. terreus 9a1        KhsdCNSVDh GYQCfPELSH kWGlYAPYFS LqDESPFPlD VPeDCHITFV
A. terreus cbs        NhsdCtSVDr GYQCfPELSH kWGlYAPYFS LqDESPFPlD VPdDCHITFV
A. niger var. awamori NqsTCDTVDq GYQCfSEtSH LWGQYAPFFS LANESAISPD VPaGCRVTFa
A. niger NRRL3135     NqsSCDTVDq GYQCfSEtSH LWGQYAPFFS LANESvISPE VPaGCRVTFa
A. fumigatus 13073    GSkSCDTVDl GYQCsPAtSH LWGQYSPFFS LEDElSVSSK LPkDCRITLV
A. fumigatus 32722    GSkSCDTVDl GYQCsPAtSH LWGQYSPFFS LEDElSVSSK LPkDCRITLV
A. fumigatus 58128    GSkSCDTVDl GYQCsPAtSH LWGQYSPFFS LEDElSVSSK LPkDCRITLV
A. fumigatus 26906    GSkSCDTVDl GYQCsPAtSH LWGQYSPFFS LEDElSVSSK LPkDCRITLV
A. fumigatus 32239    GSkACDTVEl GYQCsPGtSH LWGQYSPFFS LEDElSVSSD LPkDCRVTFV
E. nidulans           QNHSCNTaDG GYQCfPNVSH VWGQYSPYFS IEQESAISeD VPhGCeVTFV
T. thermophilus       DSHSCNTVEG GYQCrPEISH sWGQYSPFFS LADQSEISPD VPqNCKITFV
T. lanuginosa         ---------- ----nvDIAR hWGQYSPFFS LAEvSEISPA VPkGCRVeFV
M. thermophila        ESRPCDTpDl GFQCgTAISH FWGQYSPYFS VPsElDaS.. IPdDCeVTFa
Basidio               xSxPxrxtAA qLPipxQxqx xWSPYSPYFP VAxyxA.... pPaGCQIxqV

            Consensus NSHSCDTVDG GYQC-PEISH LWGQYSPFFS LADESAISPD VP-GCRVTFV
                Fcp10 NSHSCDTVDG GYQCFPEISH LWGQYSPFFS LADESAISPD VPKGCRVTFV


                      51                                                 100
A. terreus 9a1        QVLARHGARs PThSKTKaYA AtIaAIQKSA TaFpGKYAFL QSYNYSLDSE
A. terreus cbs        QVLARHGARs PTdSKTKaYA AtIaAIQKNA TaLpGKYAFL KSYNYSMGSE
A. niger var. awamori QVLSRHGARY PTeSKGKKYS ALIeEIQQNv TtFDGKYAFL KTYNYSLGAD
A. niger NRRL3135     QVLSRHGARY PTdSKGKKYS ALIeEIQQNA TtFDGKYAFL KTYNYSLGAD
A. fumigatus 13073    QVLSRHGARY PTSSKSKKYk kLVtAIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus 32722    QVLSRHGARY PTSSKSKKYk kLVtAIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus 58128    QVLSRHGARY PTSSKSKKYk kLVtAIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus 26906    QVLSRHGARY PTSSKSKKYk kLVtAIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus 32239    QVLSRHGARY PTASKSKKYk kLVtAIQKNA TeFKGKFAFL ETYNYTLGAD
E. nidulans           QVLSRHGARY PTeSKSKaYS GLIeAIQKNA TsFwGQYAFL ESYNYTLGAD
T. thermophilus       QLLSRHGARY PTSSKTElYS qLIsrIQKtA TaYKGyYAFL KdYrYqLGAN
T. lanuginosa         QVLSRHGARY PTAhKSEvYA ELLqrIQDtA TeFKGDFAFL RdYaYhLGAD
M. thermophila        QVLSRHGARa PTlkRAasYv DLIdrIHhGA isYgPgYEFL RTYDYTLGAD
Basidio               NIIqRHGARF PTSGaAtRiq AaVakLQsax xxtDPKLDFL xnxtYxLGxD

            Consensus QVLSRHGARY PTSSKSKKYS ALI-AIQKNA T-FKGKYAFL KTYNYTLGAD
                Fcp10 QVLSRHGARY PTSSKSKKYS ALIEAIQKNA TAFKGKYAFL KTYNYTLGAD


                      101                                                150
A. terreus 9a1        ELTPFGrNQL rDlGaQFYeR YNAL.TRhIn PFVRATDAsR VhESAEKFVE
A. terreus cbs        NLTPFGrNQL qDlGaQFYRR YDTL.TRhIn PFVRAADSsR VhESAEKFVE
A. niger var. awamori DLTPFGEQEL VNSGIKFYQR YESL.TRnII PFIRSSGSsR VIASGEKFIE
A. niger NRRL3135     DLTPFGEQEL VNSGIKFYQR YESL.TRnIV PFIRSSGSsR VIASGKKFIE
A. fumigatus 13073    DLTPFGEQQL VNSGIKFYQR YKAL.ARsVV PFIRASGSDR VIASGEKFIE
A. fumigatus 32722    DLTPFGEQQL VNSGIKFYQR YKAL.ARsVV PFIRASGSDR VIASGEKFIE
A. fumigatus 58128    DLTPFGEQQL VNSGIKFYQR YKAL.ARsVV PFIRASGSDR VIASGEKFIE
A. fumigatus 26906    DLTAFGEQQL VNSGIKFYQR YKAL.ARsVV PFIRASGSDR VIASGEKFIE
A. fumigatus 32239    DLTPFGEQQM VNSGIKFYQK YKAL.AgsVV PFIRSSGSDR VIASGEKFIE
E. nidulans           DLTiFGENQM VDSGaKFYRR YKnL.ARknt PFIRASGSDR VVASAEKFIN
T. thermophilus       DLTPFGENQM IQlGIKFYnH YKSL.ARnaV PFVRCSGSDR VIASGrlFIE
T. lanuginosa         NLTRFGEEQM MESGrQFYHR YREq.AReIV PFVRAAGSAR VIASAEfFnr
M. thermophila        ELTRtGQQQM VNSGIKFYRR YRAL.ARksI PFVRTAGqDR VVhSAENFtQ
Basidio               DLvPFGAxQs sQAGqEaFtR YsxLvSxdnL PFVRASGSDR VVDSAtNWtA

            Consensus DLTPFGEQQM VNSGIKFYRR YKAL-AR-IV PFVRASGSDR VIASAEKFIE
```

Fig. 7A

```
                      Fcp10  DLTPFGEQQM VNSGIKFYRR YKAL.ARKIV PFVRASGSDR VIASAEKFIE

                             151                                            200
A. terreus 9a1               GFQTARqDDh hAnphQPSPr VDVaIPEGsA YNNTLEHSLC TAFEs...St
A. terreus cbs               GFQNARqGDP hAnphQPSPr VDVVIPEGtA YNNTLEHSIC TAFEa...St
A. niger var. awamori        GFQSTKLkDP rAqpgQSSPk IDVVISEAsS sNNTLDpGtC TvFEd...SE
A. niger NRRL3135            GFQSTKLkDP rAqpgQSSPk IDVVISEAsS sNNTLDpGtC TvFEd...SE
A. fumigatus 13073           GFQqAKLADP gAt.nRAAPa ISVIIPESeT FNNTLDHGVC TkFEa...SQ
A. fumigatus 32722           GFQqAKLADP gAt.nRAAPa ISVIIPESeT FNNTLDHGVC TkFEa...SQ
A. fumigatus 58128           GFQqAKLADP gAt.nRAAPa ISVIIPESeT FNNTLDHGVC TkFEa...SQ
A. fumigatus 26906           GFQqAKLADP gAt.nRAAPa ISVIIPESeT FNNTLDHGVC TkFEa...SQ
A. fumigatus 32239           GFQqANVADP gAt.nRAAPV ISVIIPESeT YNNTLDHSVC TnFEa...SE
E. nidulans                  GFRkAQLhDh g.s.gQATPV VNVIIPEidG FNNTLDHStC vSFEn...dE
T. thermophilus              GFQSAKVlDP hSdkhDAPPt INVIIeEGpS YNNTLDtGsC PvFEd...Ss
T. lanuginosa                GFQdAKdrDP rSnkdQAePV INVIISEEtG sNNTLDgltC PAaEe...Ap
M. thermophila               GFHSAlLADR gStvrPTlPy dmVVIPETaG aNNTLHNDLC TAFEegPySt
Basidio                      GFaxA..... ..sxntxxPx LxVILSExg. .NDTLDDNMC ......PxAG

                  Consensus  GFQSAKLADP -A---QASPV INVIIPEG-G YNNTLDHGLC TAFE--P-SE
                      Fcp10  GFQSAKLADP GANPHQASPV INVIIPEGAG YNNTLDHGLC TAFEE...SE

                             201                                            250
A. terreus 9a1               VGDDavANFT AVFAPAIaqR LEAdLPGVQL StDDVVNLMA MCPFETVSlT
A. terreus cbs               VGDAaADNFT AVFAPAIakR LEAdLPGVQL SADDVVNLMA MCPFETVSlT
A. niger var. awamori        LADtVEANFT AtFAPSIRqR LEndLSGVtL TDtEVtyLMD MCSFDTIStS
A. niger NRRL3135            LADtVEANFT AtFvPSIRqR LEndLSGVtL TDtEVtyLMD MCSFDTIStS
A. fumigatus 13073           LGDEVAANFT ALFAPdIRAR aEkhLPGVtL TDEDVVSLMD MCSFDTVArT
A. fumigatus 32722           LGDEVAANFT ALFAPdIRAR aEkhLPGVtL TDEDVVSLMD MCSFDTVArT
A. fumigatus 58128           LGDEVAANFT ALFAPdIRAR aEkhLPGVtL TDEDVVSLMD MCSFDTVArT
A. fumigatus 26906           LGDEVAANFT ALFAPdIRAR aKkhLPGVtL TDEDVVSLMD MCSFDTVArT
A. fumigatus 32239           LGDEVEANFT ALFAPAIRAR IEkhLPGVQL TDDDVVSLMD MCSFDTVArT
E. nidulans                  rADEIEANFT AIMGPPIRkR LEndLPGIKL TNENVIyLMD MCSFDTMArT
T. thermophilus              gGHDaQEKFA kqFAPAIlEK IKDhLPGVDL AvsDVpyLMD LCPFETLArn
T. lanuginosa                .DptqpAEFl qVFGPRVlkK ItkhMPGVNL TlEDVplFMD LCPFDTVGsd
M. thermophila               IGDDaQDtYl StFAGPItAR VNAnLPGaNL TDADtVaLMD LCPFETVAsS
Basidio                      dSDpqxnxWl AVFAPPItAR LNAaaPGaNL TDxDaxNLxx LCPFETVS..

                  Consensus  LGDDVEANFT AVFAPPIRAR LEA-LPGVNL TDEDVVNLMD MCPFDTVA-T
                      Fcp10  LGDDVEANFT AVFAPPIRAR LEAHLPGVNL TDEDVVNLMD MCPFDTVART

                             251                                            300
A. terreus 9a1               dD..Aht... ......LSPF CDLFTa..tE WtQYNYLlSL dKYYGYGGGN
A. terreus cbs               dD..Aht... ......LSPF CDLFTa..aE WtQYNYLlSL dKYYGYGGGN
A. niger var. awamori        Tv..DTK... ......LSPF CDLFTH..dE WiHYDYLQSL kKYYGHGAGN
A. niger NRRL3135            Tv..DTK... ......LSPF CDLFTH..dE WiNYDYLQSL kKYYGHGAGN
A. fumigatus 13073           SD..ASQ... ......LSPF CQLFTH..nE WkKYNYLQSL gKYYGYGAGN
A. fumigatus 32722           SD..ASQ... ......LSPF CQLFTH..nE WkKYNYLQSL gKYYGYGAGN
A. fumigatus 58128           SD..ASQ... ......LSPF CQLFTH..nE WkKYNYLQSL gKYYGYGAGN
A. fumigatus 26906           SD..ASQ... ......LSPF CQLFTH..nE WkKYNYLQSL gKYYGYGAGN
A. fumigatus 32239           AD..ASE... ......LSPF CAIFTH..nE WkKYDYLQSL gKYYGYGAGN
E. nidulans                  AH..GTE... ......LSPF CAIFTE..kE WlQYDYLQSL sKYYGYGAGS
T. thermophilus              ht..DT.... ......LSPF CALsTQ..eE WqaYDYYQSL gKYYGnGGGN
T. lanuginosa                PvlfPrQ... ......LSPF CHLFTa..dD WmaYDYYyTL dKYYSHGGGS
M. thermophila               SsdpATadag ggngrpLSPF CrLFSE..sE WraYDYLQSV gKWYGYGPGN
Basidio                      .......... ...xexxSxF CDLFexxpeE FxaFxYxgdL dKFYGtGyGQ

                  Consensus  SD--ATQ--- ------LSPF CDLFTH---E W-QYDYLQSL -KYYGYGAGN
                      Fcp10  SD..ATQ... ......LSPF CDLFTH..DE WIQYDYLQSL GKYYGYGAGN

                             301                                            350
A. terreus 9a1               PLGPvQGVGW aNELMARLTR A.PVHDHTCv NNTLDASPAT FPLNATLYAD
A. terreus cbs               PLGPvQGVGW aNELIARLTR S.PVHDHTCv NNTLDANPAT FPLNATLYAD
A. niger var. awamori        PLGPTQGVGY aNELIARLTH S.PVHDDTSS NHTLDSNPAT FPLNSTLYAD
A. niger NRRL3135            PLGPTQGVGY aNELIARLTH S.PVHDDTSS NHTLDSSPAT FPLNSTLYAD
```

Fig. 7B

```
A. fumigatus 13073      PLGPAQGIGF tNELIARLTR S.PVQDHTST NsTLvSNPAT FPLNATMYvD
A. fumigatus 32722      PLGPAQGIGF tNELIARLTR S.PVQDHTST NsTLvSNPAT FPLNATMYvD
A. fumigatus 58128      PLGPAQGIGF tNELIARLTR S.PVQDHTST NsTLvSNPAT FPLNATMYvD
A. fumigatus 26906      PLGPAQGIGF tNELIARLTR S.PVQDHTST NsTLvSNPAT FPLNATMYvD
A. fumigatus 32239      PLGPAQGIGF tNELIARLTN S.PVQDHTST NsTLDSDPAT FPLNATIYvD
E. nidulans             PLGPAQGIGF tNELIARLTQ S.PVQDNTST NHTLDSNPAT FPLDrkLYAD
T. thermophilus         PLGPAQGVGF vNELIARMTH S.PVQDYTTv NHTLDSNPAT FPLNATLYAD
T. lanuginosa           AFGPSRGVGF vNELIARMTg NlPVKDHTTv NHTLDdNPET FPLDAvLYAD
M. thermophila          PLGPTQGVGF vNELLARLA. GvPVRDgTST NRTLDGDPrT FPLGrPLYAD
Basidio                 PLGPvQGVGY iNELLARLTx qa.VRDNTqT NRTLDSSPxT FPLNrTFYAD

              Consensus PLGPAQGVGF -NELIARLTH S-PVQDHTST NHTLDSNPAT FPLNATLYAD
                  Fcp10 PLGPAQGVGF VNELIARLTH S.PVQDHTST NHTLDSNPAT FPLNATLYAD

                        351                                              400
A. terreus 9a1          FSHDSnLVSI FWALGLYNGT aPLSqTSVE. .SvsQTDGYA AAWTVPFAAR
A. terreus cbs          FSHDSnLVSI FWALGLYNGT kPLSqTTVE. .ditrTDGYA AAWTVPFAAR
A. niger var. awamori   FSHDNGIISI LFALGLYNGT kPLSTTTVE. .NitQTDGFS SAWTVPFASR
A. niger NRRL3135       FSHDNGIISI LFALGLYNGT kPLSTTTVE. .NitQTDGFS SAWTVPFASR
A. fumigatus 13073      FSHDNSMVSI FFALGLYNGT ePLSrTSVE. .SaKElDGYS ASWvVPFGAR
A. fumigatus 32722      FSHDNSMVSI FFALGLYNGT gPLSrTSVE. .SaKElDGYS ASWvVPFGAR
A. fumigatus 58128      FSHDNSMVSI FFALGLYNGT ePLSrTSVE. .SaKElDGYS ASWvVPFGAR
A. fumigatus 26906      FSHDNSMVSI FFALGLYNGT ePLSrTSVE. .SaKElDGYS ASWvVPFGAR
A. fumigatus 32239      FSHDNGMIPI FFAMGLYNGT ePLSqTSeE. .StKESNGYS ASWAVPFGAR
E. nidulans             FSHDNSMISI FFAMGLYNGT qPLSmdSVE. .SiQEmDGYA ASWTVPFGAR
T. thermophilus         FSHDNTMtSI FaALGLYNGT akLSTTeIK. .SiEETDGYS AAWTVPFGGR
T. lanuginosa           FSHDNTMtGI FsAMGLYNGT kPLSTSkIQP pTgAAADGYA ASWTVPFAAR
M. thermophila          FSHDNdMMGV LgALGaYDGv pPLdkTA..R rdpEElGGYA ASWAVPFAAR
Basidio                 FSHDNqMVAI FsAMGLFNqS aPLdPSxpDP nrt.....Wv TSklVPFSAR

              Consensus FSHDNTMVSI FFALGLYNGT -PLSTTSVEP -S-EETDGYA ASWTVPFAAR
                  Fcp10 FSHDNTMVSI FFALGLYNGT KPLSTTSVE. .SIEETDGYA ASWTVPFAAR

                        401                                              450
A. terreus 9a1          AYVEMMQC.. ra........ .....EKEPL VRVLVNDRVM PLHGCPtDKL
A. terreus cbs          AYIEMMQC.. ra........ .....EKQPL VRVLVNDRVM PLHGCAVDNL
A. niger var. awamori   lYVEMMQC.. Qa........ .....EQEPL VRVLVNDRVV PLHGCPIDaL
A. niger NRRL3135       lYVEMMQC.. Qa........ .....EQEPL VRVLVNDRVV PLHGCPVDaL
A. fumigatus 13073      AYfEtMQC.. Ks........ .....EKEPL VRaLINDRVV PLHGCDVDKL
A. fumigatus 32722      AYfEtMQC.. Ks........ .....EKEPL VRaLINDRVV PLHGCDVDKL
A. fumigatus 58128      AYfEtMQC.. Ks........ .....EKESL VRaLINDRVV PLHGCDVDKL
A. fumigatus 26906      AYfEtMQC.. Ks........ .....EKEPL VRaLINDRVV PLHGCDVDKL
A. fumigatus 32239      AYfEtMQC.. Ks........ .....EKEPL VRaLINDRVV PLHGCAVDKL
E. nidulans             AYfELMQC.. E......... .....KKEPL VRVLVNDRVV PLHGCAVDKF
T. thermophilus         AYIEMMQC.. Dd........ .....sDEPV VRVLVNDRVV PLHGCEVDsL
T. lanuginosa           AYVELLRC.. Etetsseeee EG...EDEPF VRVLVNDRVV PLHGCrVDRW
M. thermophila          iYVEkMRC.. sggggggggg EGrqeKDEeM VRVLVNDRVM TLkGCGaDEr
Basidio                 mvVErLxCxx xgtxxxxxxx xxxxxxxxxx VRVLVNDaVq PLEfCGgDxd

              Consensus AYVEMMQC-- E--------- EG---EKEPL VRVLVNDRVV PLHGCGVDKL
                  Fcp10 AYVEMMQC.. EA........ .....EKEPL VRVLVNDRVV PLHGCGVDKL
```

Fig. 7C

```
                      451                               482
A. terreus 9al        GRCKrDAFVA GLSFAQAG.. GNWADCF--- --
A. terreus cbs        GRCKrDDFVE GLSFARAG.. GNWAECF--- --
A. niger var. awamori GRCtrDsFVr GLSFARSG.. GDWAECsA-- --
A. niger NRRL3135     GRCtrDsFVr GLSFARSG.. GDWAECFA-- --
A. fumigatus 13073    GRCKlNDFVK GLSWARSG.. GNWGECFS-- --
A. fumigatus 32722    GRCKlNDFVK GLSWARSG.. GNWGECFS-- --
A. fumigatus 58128    GRCKlNDFVK GLSWARSG.. GNWGECFS-- --
A. fumigatus 26906    GRCKlNDFVK GLSWARSG.. GNWGECFS-- --
A. fumigatus 32239    GRCKlKDFVK GLSWARSG.. GNSEQSFS-- --
E. nidulans           GRCtlDDWVE GLNFARSG.. GNWKtCFTl- --
T. thermophilus       GRCKrDDFVr GLSFARqG.. GNWEGCYAas e-
T. lanuginosa         GRCRrDEWIK GLTFARqG.. GHWDrCF--- --
M. thermophila        GmCtlErFIE SMAFARGN.. GKWDlCFA-- --
Basidio               GxCtlDAFVE SqxYAReDgq GDFEKCFAtp xx

            Consensus GRCK-DDFVE GLSFARSG-- GNWEECFA-- --
                Fcp10 GRCKRDDFVE GLSFARSG.. GNWEECFA.. ..
```

Fig. 7D

```
                         1                                                   50
P. involutus (phyA1)     ---------- -FPipeseqR nWSPYSPYFP LAEykA.... pPaGCQInqV
P. involutus (phyA2)     ---------- -FsipeseqR nWSPYSPYFP LAEykA.... pPaGCeInqV
T. pubescens             ---------- -LDvtRDVqQ sWSmYSPYFP aAtyvA.... pPaSCQInqV
A. pediades              ---------- -pffpPQIqD sWAaYTPYYP VqAyTP.... pPKDCKITqV
P. lycii                 ---------- -LPipAQnTs nWGPYdPFFP VEpyAA.... pPEGCtVTqV
A. terreus 9a1           KhsdCNSVDh GYQCfPELSH kWGlYAPYFS LqDESPFPlD VPEDCHITFV
A. terreus cbs           NhsdCtSVDr GYQCfPELSH kWGlYAPYFS LqDESPFPlD VPDDCHITFV
A. niger var. awamori    NqsTCDTVDq GYQCfSEtSH LWGQYAPFFS LANESAISPD VPaGCRVTFa
A. niger T213            NqsSCDTVDq GYQCfSEtSH LWGQYAPFFS LANESvISPD VPaGCRVTFa
A. niger NRRL3135        NqsSCDTVDq GYQCfSEtSH LWGQYAPFFS LANESvISPE VPaGCRVTFa
A. fumigatus ATCC13073   GSkSCDTVDl GYQCsPAtSH LWGQYSPFFS LEDElSVSSK LPKDCRITLV
A. fumigatus ATCC32722   GSkSCDTVDl GYQCsPAtSH LWGQYSPFFS LEDElSVSSK LPKDCRITLV
A. fumigatus ATCC58128   GSkSCDTVDl GYQCsPAtSH LWGQYSPFFS LEDElSVSSK LPKDCRITLV
A. fumigatus ATCC26906   GSkSCDTVDl GYQCsPAtSH LWGQYSPFFS LEDElSVSSK LPKDCRITLV
A. fumigatus ATCC32239   GSkACDTVEl GYQCsPGtSH LWGQYSPFFS LEDElSVSSD LPKDCRVTFV
E. nidulans              QNHSCNTaDg GYQCfPNVSH VWGQYSPYFS IEQESAISeD VPhGCeVTFV
T. thermophilus          DSHSCNTVEg GYQCrPEISH sWGQYSPFFS LADQSEISPD VPQNCKITFV
T. lanuginosa            ---------- ----nvDIAR hWGQYSPFFS LAEvSEISPA VPKGCRVeFV
M. thermophila           ESRPCDTpDl GFQCgTAISH FWGQYSPYFS VPsElDaS.. IPDDCeVTFa

Consensus Seq. 11        NSHSCDTVD- GYQC-PEISH LWGQYSPFFS LADESAISPD VPKGCRVTFV

                         51                                                 100
P. involutus (phyA1)     NIIqRHGARF PTSGaTtRik AgLtKLQgvq nftDAKFnFI KSFKYdLGns
P. involutus (phyA2)     NIIqRHGARF PTSGaAtRik AgLsKLQsvq nftDPKFDFI KSFtYdLGTs
T. pubescens             HIIqRHGARF PTSGaAKRiq TaVAKLKaaS nytDPlLAFV tnYtYSLGqD
A. pediades              NIIqRHGARF PTSGaGtRiq AaVKKLQsak TytDPRLDFL tnYtYTLGhD
P. lycii                 NLIqRHGARW PTSGarsRqv AaVAKIQmar PftDPKYEFL NdFvYkFGvA
A. terreus 9a1           QVLARHGARs PThSKTKaYA AtIAaIQKSA TaFpGKYAFL QSYNYSLDSE
A. terreus cbs           QVLARHGARs PTdSKTKaYA AtIAaIQKNA TaLpGKYAFL KSYNYSMGSE
A. niger var. awamori    QVLSRHGARY PTeSKGKKYS ALIEeIQQNv TtFDGKYAFL KTYNYSLGAD
A. niger T213            QVLSRHGARY PTeSKGKKYS ALIEeIQQNv TtFDGKYAFL KTYNYSLGAD
A. niger NRRL3135        QVLSRHGARY PTdSKGKKYS ALIEeIQQNA TtFDGKYAFL KTYNYSLGAD
A. fumigatus ATCC13073   QVLSRHGARY PTSSKSKKYk kLVtaIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus ATCC32722   QVLSRHGARY PTSSKSKKYk kLVtaIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus ATCC58128   QVLSRHGARY PTSSKSKKYk kLVtaIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus ATCC26906   QVLSRHGARY PTSSKSKKYk kLVtaIQaNA TdFKGKFAFL KTYNYTLGAD
A. fumigatus ATCC32239   QVLSRHGARY PTASKSKKYk kLVtaIQKNA TeFKGKFAFL ETYNYTLGAD
E. nidulans              QVLSRHGARY PTeSKSKaYS GLIEaIQKNA TsFwGQYAFL ESYNYTLGAD
T. thermophilus          QLLSRHGARY PTSSKTElYS qLIsRIQKtA TaYKGyYAFL KdYrYqLGAN
T. lanuginosa            QVLSRHGARY PTAhKSEvYA ELLQRIQDtA TeFKGDFAFL RdYaYhLGAD
M. thermophila           QVLSRHGARa PTlkRAasYv DLIDRIHhGA isYgPgYEFL RTYDYTLGAD

Consensus Seq. 11        QVLSRHGARY PTSSKSKKYS ALIERIQKNA T-FKGKYAFL KTYNYTLGAD


                  101
150
P. involutus (phyA1)     DLvPFGAaQs fDAGqEaFaR YskLvSKNnL PFIRAdGSDR VVDSAtNWtA
P. involutus (phyA2)     DLvPFGAaQs fDAGLEvFaR YskLvSsDnL PFIRSdGSDR VVDTAtNWtA
T. pubescens             sLveLGAtQs sEAGqEaFtR YsSLvSaDeL PFVRASGSDR VVATANNWtA
A. pediades              DLvPFGAlQs sQAGeEtFQR YsfLvSKEnL PFVRASSSNR VVDSAtNWtE
P. lycii                 DLlPFGANQs hQTGtDMYtR YsTLfEgGdV PFVRAAGdQR VVDSStNWtA
A. terreus 9a1           ELTPFGrNQL rDlGaQFYeR YNAL.TRHIn PFVRATDAsR VhESAEKFVE
A. terreus cbs           NLTPFGrNQL qDlGaQFYRR YDTL.TRHIn PFVRAADSsR VhESAEKFVE
A. niger var. awamori    DLTPFGEQEL VNSGIKFYQR YESL.TRNII PFIRSSGSsR VIASGEKFIE
A. niger T213            DLTPFGEQEL VNSGIKFYQR YESL.TRNII PFIRSSGSsR VIASGEKFIE
A. niger NRRL3135        DLTPFGEQEL VNSGIKFYQR YESL.TRNIV PFIRSSGSsR VIASGKKFIE
A. fumigatus ATCC13073   DLTPFGEQQL VNSGIKFYQR YKAL.ARSVV PFIRASGSDR VIASGEKFIE
A. fumigatus ATCC32722   DLTPFGEQQL VNSGIKFYQR YKAL.ARSVV PFIRASGSDR VIASGEKFIE
A. fumigatus ATCC58128   DLTPFGEQQL VNSGIKFYQR YKAL.ARSVV PFIRASGSDR VIASGEKFIE
```

Fig. 8A

| | | | | | |
|---|---|---|---|---|---|
| *A. fumigatus* ATCC26906 | DLTAFGEQQL | VNSGIKFYQR | YKAL.ARSVV | PFIRASGSDR | VIASGEKFIE |
| *A. fumigatus* ATCC32239 | DLTPFGEQQM | VNSGIKFYQK | YKAL.AgSVV | PFIRSSGSDR | VIASGEKFIE |
| *E. nidulans* | DLTiFGENQM | VDSGaKFYRR | YKnL.ARKnt | PFIRASGSDR | VVASAEKFIN |
| *T. thermophilus* | DLTPFGENQM | IQlGIKFYnH | YKSL.ARNaV | PFVRCSGSDR | VIASGrlFIE |
| *T. lanuginosa* | NLTRFGEEQM | MESGrQFYHR | YREq.AREIV | PFVRAAGSAR | VIASAEfFnr |
| *M. thermophila* | ELTRtGQQQM | VNSGIKFYRR | YRAL.ARKsI | PFVRTAGqDR | VVhSAENFtQ |
| **Consensus Seq. 11** | **DLTPFGENQM** | **VNSGIKFYRR** | **YKAL-ARNIV** | **PFVRASGSDR** | **VIASAEKFIE** |

| | 151 | | | | 200 |
|---|---|---|---|---|---|
| *P. involutus (phyA1)* | GFaSA..... | ..shNtvqPk | LNLILPQ..T | gNDTLEDNMC | PAaGD..... |
| *P. involutus (phyA2)* | GFaSA..... | ..srNaiqPk | LDLILPQ..T | gNDTLEDNMC | PAaGE..... |
| *T. pubescens* | GFalA..... | ..ssNsiTPV | LSVIISE..A | gNDTLDDNMC | PAaGD..... |
| *A. pediades* | GFsAA..... | ..shHvlNPI | LfVILSE..S | LNDTLDDAMC | PnaGs..... |
| *P. lycii* | GFgdA..... | ..sgEtvlPt | LQVVLQE..E | gNcTLcNNMC | PnevD..... |
| *A. terreus* 9a1 | GFQTARqDDh | hAnpHQPSPr | VDVaIPEGSA | YNNTLEHSLC | TAFEs...ST |
| *A. terreus* cbs | GFQNARqGDP | hAnpHQPSPr | VDVVIPEGTA | YNNTLEHSIC | TAFEA...ST |
| *A. niger* var. *awamori* | GFQSTKLkDP | rAqpgQSSPk | IDVVISEASS | sNNTLDpGtC | TvFED...Se |
| *A. niger* T213 | GFQSTKLkDP | rAqpgQSSPk | IDVVISEASS | sNNTLDpGtC | TvFED...Se |
| *A. niger* NRRL3135 | GFQSTKLkDP | rAqpgQSSPk | IDVVISEASS | sNNTLDpGtC | TvFED...Se |
| *A. fumigatus* ATCC13073 | GFQqAKLADP | gAt.NRAAPa | ISVIIPESeT | FNNTLDHGVC | TkFEA...Sq |
| *A. fumigatus* ATCC32722 | GFQqAKLADP | gAt.NRAAPa | ISVIIPESeT | FNNTLDHGVC | TkFEA...Sq |
| *A. fumigatus* ATCC58128 | GFQqAKLADP | gAt.NRAAPa | ISVIIPESeT | FNNTLDHGVC | TkFEA...Sq |
| *A. fumigatus* ATCC26906 | GFQqAKLADP | gAt.NRAAPa | ISVIIPESeT | FNNTLDHGVC | TkFEA...Sq |
| *A. fumigatus* ATCC32239 | GFQqANVADP | gAt.NRAAPV | ISVIIPESeT | YNNTLDHSVC | TnFEA...Se |
| *E. nidulans* | GFRkAQLhDh | g.s.gQATPV | VNVIIPEidG | FNNTLDHStC | vSFEN...de |
| *T. thermophilus* | GFQSAKVlDP | hSdkHDAPPt | INVIIeEGPS | YNNTLDtGsC | PvFED...SS |
| *T. lanuginosa* | GFQdAKdrDP | rSnkDQAePV | INVIISEETG | sNNTLDgltC | PAaEE...AP |
| *M. thermophila* | GFHSAlLADR | gStvRPTlPy | dmVVIPETAG | aNNTLHNDLC | TAFEEgpyST |
| **Consensus Seq. 11** | **GFQSAKLADP** | **-A--HQASPV** | **INVIIPEGSG** | **YNNTLDHGLC** | **TAFED---ST** |

| | 201 | | | | 250 |
|---|---|---|---|---|---|
| *P. involutus (phyA1)* | .SDpqvnaWl | AVafPSItAR | LNAaaPSVNL | TDtDafNLVs | LCAFlTVSK. |
| *P. involutus (phyA2)* | .SDpqvDaWl | AsafPSVtAQ | LNAaaPGaNL | TDADafNLVs | LCPFmTVSK. |
| *T. pubescens* | .SDpqvnQWl | AqFAPPMtAR | LNAgaPGaNL | TDtDtyNLLt | LCPFETVAt. |
| *A. pediades* | .SDpqtGiWT | SIYGTPIanR | LNqqaPGaNI | TAADVsNLIp | LCAFETIvK. |
| *P. lycii* | .GDESt.tWl | GVFAPnItAR | LNAaaPSaNL | SDsDaLtLMD | MCPFDTLSs. |
| *A. terreus* 9a1 | VGDDAvANFT | AVFAPAIaqR | LEAdLPGVQL | StDDVVNLMA | MCPFETVSlT |
| *A. terreus* cbs | VGDAAADNFT | AVFAPAIakR | LEAdLPGVQL | SADDVVNLMA | MCPFETVSlT |
| *A. niger* var. *awamori* | LADtvEANFT | AtFAPSIRqR | LEndLSGVtL | TDtEVtyLMD | MCSFDTIStS |
| *A. niger* T213 | LADtvEANFT | AtFAPSIRqR | LEndLSGVtL | TDtEVtyLMD | MCSFDTIStS |
| *A. niger* NRRL3135 | LADtvEANFT | AtFvPSIRqR | LEndLSGVtL | TDtEVtyLMD | MCSFDTIStS |
| *A. fumigatus* ATCC13073 | LGDEvAANFT | ALFAPdIRAR | aEkhLPGVtL | TDEDVVSLMD | MCSFDTVART |
| *A. fumigatus* ATCC32722 | LGDEvAANFT | ALFAPdIRAR | aEkhLPGVtL | TDEDVVSLMD | MCSFDTVART |
| *A. fumigatus* ATCC58128 | LGDEvAANFT | ALFAPdIRAR | aEkhLPGVtL | TDEDVVSLMD | MCSFDTVART |
| *A. fumigatus* ATCC26906 | LGDEvAANFT | ALFAPdIRAR | aKkhLPGVtL | TDEDVVSLMD | MCSFDTVART |
| *A. fumigatus* ATCC32239 | LGDEvEANFT | ALFAPAIRAR | IEkhLPGVQL | TDDDVVSLMD | MCSFDTVART |
| *E. nidulans* | rADEiEANFT | AIMGPPIRkR | LEndLPGIKL | TNENVIyLMD | MCSFDTMART |
| *T. thermophilus* | gGHDAQEKFA | kqFAPAIlEK | IKDhLPGVDL | AvsDVpyLMD | LCPFETLARn |
| *T. lanuginosa* | .DptqpAEFl | qVFGPRVlkK | ItkhMPGVNL | TlEDVplFMD | LCPFDTVGsd |
| *M. thermophila* | IGDDAQDtYl | StFAGPItAR | VNAnLPGaNL | TDADtVaLMD | LCPFETVAsS |
| **Consensus Seq. 11** | **LGDDAEANFT** | **AVFAPPIRAR** | **LEA-LPGVNL** | **TDEDVVNLMD** | **MCPFDTVART** |

| | 251 | | | | 300 |
|---|---|---|---|---|---|
| *P. involutus (phyA1)* | .......... | ....ekkSdF | CtLFegiPGs | FeaFAYggdL | dKFYGtGyGQ |
| *P. involutus (phyA2)* | .......... | ....eqkSdF | CtLFegiPGs | FeaFAYagdL | dKFYGtGyGQ |
| *T. pubescens* | .......... | ....errSeF | CDIYeelqAE | .daFAYnadL | dKFYGtGyGQ |
| *A. pediades* | .......... | ....etpSPF | CNLF..TPEE | FaQFEYFgdL | dKFYGtGyGQ |
| *P. lycii* | .......... | ....gnaSPF | CDLF..TAEE | YvsYEYYydL | dKYYGtGPGN |
| *A. terreus* 9a1 | dD..Aht... | ......LSPF | CDLF..TAtE | WtQYNYLlSL | dKYYGYGGGN |
| *A. terreus* cbs | dD..Aht... | ......LSPF | CDLF..TAAE | WtQYNYLlSL | dKYYGYGGGN |
| *A. niger* var. *awamori* | Tv..DTK... | ......LSPF | CDLF..ThDE | WiHYDYLQSL | kKYYGHGAGN |

Fig. 8B

```
A. niger T213               Tv..DTK... ......LSPF CDLF..ThDE WiHYDYLRSL kKYYGHGAGN
A. niger NRRL3135           Tv..DTK... ......LSPF CDLF..ThDE WiNYDYLQSL kKYYGHGAGN
A. fumigatus ATCC13073      SD..ASQ... ......LSPF CQLF..ThNE WkKYNYLQSL gKYYGYGAGN
A. fumigatus ATCC32722      SD..ASQ... ......LSPF CQLF..ThNE WkKYNYLQSL gKYYGYGAGN
A. fumigatus ATCC58128      SD..ASQ... ......LSPF CQLF..ThNE WkKYNYLQSL gKYYGYGAGN
A. fumigatus ATCC26906      SD..ASQ... ......LSPF CQLF..ThNE WkKYNYLQSL gKYYGYGAGN
A. fumigatus ATCC32239      AD..ASE... ......LSPF CAIF..ThNE WkKYDYLQSL gKYYGYGAGN
E. nidulans                 AH..GTE... ......LSPF CAIF..TEKE WlQYDYLQSL sKYYGYGAGS
T. thermophilus             ht..DT.... ......LSPF CALs..TqEE WqaYDYYQSL gKYYGnGGGN
T. lanuginosa               PvlfPrQ... ......LSPF CHLF..TADD WmaYDYYyTL dKYYSHGGGS
M. thermophila              SsdpATadag ggngrpLSPF CrLF..SEsE WraYDYLQSV gKWYGYGPGN

Consensus Seq. 11           SD--ATQ--- ------LSPF CDLF--TADE W-QYDYLQSL -KYYGYGAGN

                            301                                              350
P. involutus (phyA1)        eLGPvQGVGY vNELIARLTN S.AVRDNTqT NRTLDASPvT FPLNkTFYAD
P. involutus (phyA2)        ALGPvQGVGY iNELLARLTN S.AVNDNTqT NRTLDAaPDT FPLNkTMYAD
T. pubescens                PLGPvQGVGY iNELIARLTa q.nVsDHTqT NsTLDSSPET FPLNrTLYAD
A. pediades                 PLGPvQGVGY iNELLARLTE m.PVRDNTqT NRTLDSSPlT FPLDrSIYAD
P. lycii                    ALGPvQGVGY vNELLARLTg q.AVRDETqT NRTLDSDPAT FPLNrTFYAD
A. terreus 9a1              PLGPvQGVGW aNELMARLTR A.PVHDHTCv NNTLDASPAT FPLNATLYAD
A. terreus cbs              PLGPvQGVGW aNELIARLTR S.PVHDHTCv NNTLDANPAT FPLNATLYAD
A. niger var. awamori       PLGPTQGVGY aNELIARLTH S.PVHDDTSS NHTLDSNPAT FPLNSTLYAD
A. niger T213               PLGPTQGVGY aNELIARLTH S.PVHDDTSS NHTLDSNPAT FPLNSTLYAD
A. niger NRRL3135           PLGPTQGVGY aNELIARLTH S.PVHDDTSS NHTLDSSPAT FPLNSTLYAD
A. fumigatus ATCC13073      PLGPAQGIGF tNELIARLTR S.PVQDHTST NsTLvSNPAT FPLNATMYvD
A. fumigatus ATCC32722      PLGPAQGIGF tNELIARLTR S.PVQDHTST NsTLvSNPAT FPLNATMYvD
A. fumigatus ATCC58128      PLGPAQGIGF tNELIARLTR S.PVQDHTST NsTLvSNPAT FPLNATMYvD
A. fumigatus ATCC26906      PLGPAQGIGF tNELIARLTR S.PVQDHTST NsTLvSNPAT FPLNATMYvD
A. fumigatus ATCC32239      PLGPAQGIGF tNELIARLTN S.PVQDHTST NsTLDSDPAT FPLNATIYvD
E. nidulans                 PLGPAQGIGF tNELIARLTQ S.PVQDNTST NHTLDSNPAT FPLDrkLYAD
T. thermophilus             PLGPAQGVGF vNELIARMTH S.PVQDYTTv NHTLDSNPAT FPLNATLYAD
T. lanuginosa               AFGPSRGVGF vNELIARMTg NlPVKDHTTv NHTLDdNPET FPLDAvLYAD
M. thermophila              PLGPTQGVGF vNELLARLA. GvPVRDgTST NRTLDGDPrT FPLGrPLYAD

Consensus Seq. 11           PLGPAQGVGF -NELIARLTH S-PVQDHTST NHTLDSNPAT FPLNATLYAD

                            351                                              400
P. involutus (phyA1)        FSHDNlMVAV FsAMGLFrqP aPLSTSvpNP wrt.....Wr TSSlVPFSGR
P. involutus (phyA2)        FSHDNlMVAV FsAMGLFrqS aPLSTSTpDP nrt.....Wl TSSvVPFSAR
T. pubescens                FSHDNqMVAI FsAMGLFNqS aPLdPTTpDP art.....Fl vkkiVPFSAR
A. pediades                 LSHDNqMIAI FsAMGLFNqS sPLdPSfpNP krt.....Wv TSRltPFSAR
P. lycii                    FSHDNTMVPI FaALGLFNAT a.LdPlkpDe nrl.....Wv DSklVPFSGH
A. terreus 9a1              FSHDSnLVSI FWALGLYNGT aPLSqTSVES Vs..QTDGYA AAWTVPFAAR
A. terreus cbs              FSHDSnLVSI FWALGLYNGT KPLSqTTVEd It..rTDGYA AAWTVPFAAR
A. niger var. awamori       FSHDNGIISI LFALGLYNGT KPLSTTTVEN It..QTDGFS SAWTVPFASR
A. niger T213               FSHDNGIISI LFALGLYNGT KPLSTTTVEN It..QTDGFS SAWTVPFASR
A. niger NRRL3135           FSHDNGIISI LFALGLYNGT KPLSTTTVEN It..QTDGFS SAWTVPFASR
A. fumigatus ATCC13073      FSHDNSMVSI FFALGLYNGT EPLSrTSVES ak..ElDGYS ASWvVPFGAR
A. fumigatus ATCC32722      FSHDNSMVSI FFALGLYNGT gPLSrTSVES ak..ElDGYS ASWvVPFGAR
A. fumigatus ATCC58128      FSHDNSMVSI FFALGLYNGT EPLSrTSVES ak..ElDGYS ASWvVPFGAR
A. fumigatus ATCC26906      FSHDNSMVSI FFALGLYNGT EPLSrTSVES ak..ElDGYS ASWvVPFGAR
A. fumigatus ATCC32239      FSHDNGMIPI FFAMGLYNGT EPLSqTSeES tk..ESNGYS ASWAVPFGAR
E. nidulans                 FSHDNSMISI FFAMGLYNGT QPLSmdSVES Iq..EmDGYA ASWTVPFGAR
T. thermophilus             FSHDNTMtSI FaALGLYNGT akLSTTeIKS Ie..ETDGYS AAWTVPFGGR
T. lanuginosa               FSHDNTMtGI FsAMGLYNGT KPLSTSkIQP ptgaAADGYA ASWTVPFAAR
M. thermophila              FSHDNdMMGV LgALGaYDGv pPLdkTArrd ..peElGGYA ASWAVPFAAR

Consensus Seq. 11           FSHDNTMVSI FFALGLYNGT KPLSTTSVES I---ETDGYA ASWTVPFAAR

                            401                                              450
P. involutus (phyA1)        mvVErLsC.. fGt....... ........Tk VRVLVQDQVq PLEfCGgDRn
P. involutus (phyA2)        maVErLsC.. AGt....... ........Tk VRVLVQDQVq PLEfCGgDQd
T. pubescens                mvVErLDC.. GGa....... ........Qs VRLLVNDaVq PLafCGaDts
```

Fig. 8C

| | | | | | |
|---|---|---|---|---|---|
| A. *pediades* | mvtErLlCQr | DGtGsGGpsr | imrNgnvQTF | VRILVNDaLq | PLkfCGgDmd |
| P. *lycii* | mtVEkLaC.. | .......... | .....sgKea | VRVLVNDaVq | PLEfCGg.vd |
| A. *terreus* 9a1 | AYVEMMQCrA | .......... | ..EK...EPL | VRVLVNDRVM | PLHGCPtDKL |
| A. *terreus* cbs | AYIEMMQCrA | .......... | ..EK...QPL | VRVLVNDRVM | PLHGCAVDNL |
| A. *niger* var. *awamori* | lYVEMMQCQA | .......... | ..EQ...EPL | VRVLVNDRVV | PLHGCPIDaL |
| A. *niger* T213 | lYVEMMQCQA | .......... | ..EQ...EPL | VRVLVNDRVV | PLHGCPIDaL |
| A. *niger* NRRL3135 | lYVEMMQCQA | .......... | ..EQ...EPL | VRVLVNDRVV | PLHGCPVDaL |
| A. *fumigatus* ATCC13073 | AYfEtMQCKS | .......... | ..EK...EPL | VRaLINDRVV | PLHGCDVDKL |
| A. *fumigatus* ATCC32722 | AYfEtMQCKS | .......... | ..EK...EPL | VRaLINDRVV | PLHGCDVDKL |
| A. *fumigatus* ATCC58128 | AYfEtMQCKS | .......... | ..EK...ESL | VRaLINDRVV | PLHGCDVDKL |
| A. *fumigatus* ATCC26906 | AYfEtMQCKS | .......... | ..EK...EPL | VRaLINDRVV | PLHGCDVDKL |
| A. *fumigatus* ATCC32239 | AYfEtMQCKS | .......... | ..EK...EPL | VRaLINDRVV | PLHGCAVDKL |
| E. *nidulans* | AYfELMQCE. | .......... | ..KK...EPL | VRVLVNDRVV | PLHGCAVDKF |
| T. *thermophilus* | AYIEMMQCDD | .......... | ..sD...EPV | VRVLVNDRVV | PLHGCEVDsL |
| T. *lanuginosa* | AYVELLRCET | ETsSeEEeEG | ..ED...EPF | VRVLVNDRVV | PLHGCrVDRW |
| M. *thermophila* | iYVEkMRCsG | GGgGgGGgEG | ..rQekdEeM | VRVLVNDRVM | TLkGCGaDEr |
| | | | | | |
| Consensus Seq. 11 | AYVEMMQCEA | GG-G-GG-EG | --EK---EPL | VRVLVNDRVV | PLHGCGVDKL |

| | 451 | | 482 | |
|---|---|---|---|---|
| P. *involutus* (*phyA1*) | GlCtLAKFVE | SqTFARSDga | GDFEKCFAts | a~ |
| P. *involutus* (*phyA2*) | GlCaLDKFVE | SqAYARSGga | GDFEKCLAtt | v~ |
| T. *pubescens* | GvCtLDAFVE | SqAYARNDge | GDFEKCFAt~ | -- |
| A. *pediades* | SlCtLEAFVE | SqkYAReDgq | GDFEKCFD-- | -- |
| P. *lycii* | GvCELsAFVE | SqTYAReNgq | GDFAKCgfvp | se |
| A. *terreus* 9a1 | GRCKrDAFVA | GLSFAQAG.. | GNWADCF--- | -- |
| A. *terreus* cbs | GRCKrDDFVE | GLSFARAG.. | GNWAECF--- | -- |
| A. *niger* var. *awamori* | GRCtrDsFVr | GLSFARSG.. | GDWAECsA-- | -- |
| A. *niger* T213 | GRCtrDsFVr | GLSFARSG.. | GDWAECFA-- | -- |
| A. *niger* NRRL3135 | GRCtrDsFVr | GLSFARSG.. | GDWAECFA-- | -- |
| A. *fumigatus* ATCC13073 | GRCKLNDFVK | GLSWARSG.. | GNWGECFS-- | -- |
| A. *fumigatus* ATCC32722 | GRCKLNDFVK | GLSWARSG.. | GNWGECFS-- | -- |
| A. *fumigatus* ATCC58128 | GRCKLNDFVK | GLSWARSG.. | GNWGECFS-- | -- |
| A. *fumigatus* ATCC26906 | GRCKLNDFVK | GLSWARSG.. | GNWGECFS-- | -- |
| A. *fumigatus* ATCC32239 | GRCKLKDFVK | GLSWARSG.. | GNSEQSFS-- | -- |
| E. *nidulans* | GRCtLDDWVE | GLNFARSG.. | GNWktCFTl- | -- |
| T. *thermophilus* | GRCKrDDFVr | GLSFARqG.. | GNWEGCYAas | e~ |
| T. *lanuginosa* | GRCRrDEWIK | GLTFARqG.. | GHWDrCF--- | -- |
| M. *thermophila* | GmCtLErFIE | SMAFARGN.. | GKWDlCFA-- | -- |
| | | | | |
| Consensus Seq. 11 | GRCKLDDFVE | GLSFARSG-- | GNWAECFA-- | -- |

Fig. 8D

```
      M  G  V  F  V  V  L  L  S  I  A  T  L  F  G  S  T  S  G  T     20
      ATGGGCGTGTTCGTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCCACATCCGGTACC
    1 ---+---------+---------+---------+---------+---------+------   60
      TACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGTGTAGGCCATGG

      A  L  G  P  R  G  N  S  H  S  C  D  T  V  D  G  G  Y  Q  C     40
      GCCTTGGGTCCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTGGTTACCAATGT
   61 ---+---------+---------+---------+---------+---------+------  120
      CGGAACCCAGGAGCACCATTAAGAGTGAGAACACTGTGACAACTGCCACCAATGGTTACA

      F  P  E  I  S  H  L  W  G  T  Y  S  P  Y  F  S  L  A  D  E     60
      TTCCCAGAAATTTCTCACTTGTGGGGTACCTACTCTCCATACTTCTCTTTGGCAGACGAA
  121 ---+---------+---------+---------+---------+---------+------  180
      AAGGGTCTTTAAAGAGTGAACACCCCATGGATGAGAGGTATGAAGAGAAACCGTCTGCTT

      S  A  I  S  P  D  V  P  D  D  C  R  V  T  F  V  Q  V  L  S     80
      TCTGCTATTTCTCCAGACGTTCCAGACGACTGTAGAGTTACTTTCGTTCAAGTTTTGTCT
  187 ---+---------+---------+---------+---------+---------+------  240
      AGACGATAAAGAGGTCTGCAAGGTCTGCTGACATCTCAATGAAAGCAAGTTCAAAACAGA

      R  H  G  A  R  Y  P  T  S  S  A  S  K  A  Y  S  A  L  I  E    100
      AGACACGGTGCTAGATACCCAACTTCTTCTGCGTCTAAGGCTTACTCTGCTTTGATTGAA
  241 ---+---------+---------+---------+---------+---------+------  300
      TCTGTGCCACGATCTATGGGTTGAAGAAGACGCAGATTCCGAATGAGACGAAACTAACTT

      A  I  Q  K  N  A  T  A  F  K  G  K  Y  A  F  L  K  T  Y  N    120
      GCTATTCAAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGAAGACTTACAAC
  301 ---+---------+---------+---------+---------+---------+------  360
      CGATAAGTTTTCTTGCGATGACGAAAGTTCCCATTCATGCGAAAGAACTTCTGAATGTTG

      Y  T  L  G  A  D  D  L  T  P  F  G  E  N  Q  M  V  N  S  G    140
      TACACTTTGGGTGCTGACGACTTGACTCCATTCGGTGAAAACCAAATGGTTAACTCTGGT
  361 ---+---------+---------+---------+---------+---------+------  420
      ATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACCAATTGAGACCA

      I  K  F  Y  R  R  Y  K  A  L  A  R  K  I  V  P  F  I  R  A    160
      ATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCATTCATTAGAGCT
  421 ---+---------+---------+---------+---------+---------+------  480
      TAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTAAGTAATCTCGA

      S  G  S  D  R  V  I  A  S  A  E  K  F  I  E  G  F  Q  S  A    180
      TCTGGTTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTTTCCAATCTGCT
  481 ---+---------+---------+---------+---------+---------+------  540
      AGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAAAGGTTAGACGA

      K  L  A  D  P  G  S  Q  P  H  Q  A  S  P  V  I  N  V  I  I    200
      AAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTAACGTGATCATT
  541 ---+---------+---------+---------+---------+---------+------  600
      TTCAACCGACTGGGTCCAAGAGTTGGTGTGGTTCGAAGAGGTCAATAATTGCACTAGTAA

      P  E  G  S  G  Y  N  N  T  L  D  H  G  T  C  T  A  F  E  D    220
      CCAGAAGGATCCGGTTACAACAACACTTTGGACCACGGTACTTGTACTGCTTTCGAAGAC
  601 ---+---------+---------+---------+---------+---------+------  660
      GGTCTTCCTAGGCCAATGTTGTTGTGAAACCTGGTGCCATGAACATGACGAAAGCTTCTG
```

Fig. 9A

```
      S  E  L  G  D  D  V  E  A  N  F  T  A  L  F  A  P  A  I  R    240
      TCTGAATTAGGTGACGACGTTGAAGCTAACTTCACTGCTTTGTTCGCTCCAGCTATTAGA
 661  ---+---------+---------+---------+---------+---------+------   720
      AGACTTAATCCACTGCTGCAACTTCGATTGAAGTGACGAAACAAGCGAGGTCGATAATCT

      A  R  L  E  A  D  L  P  G  V  T  L  T  D  E  D  V  V  Y  L    260
      GCTAGATTGGAAGCTGACTTGCCAGGTGTTACTTTGACTGACGAAGACGTTGTTTACTTG
 721  ---+---------+---------+---------+---------+---------+------   780
      CGATCTAACCTTCGACTGAACGGTCCACAATGAAACTGACTGCTTCTGCAACAAATGAAC

      M  D  M  C  P  F  D  T  V  A  R  T  S  D  A  T  E  L  S  P    280
      ATGGACATGTGTCCATTCGACACTGTCGCTAGAACTTCTGACGCTACTGAATTGTCTCCA
 781  ---+---------+---------+---------+---------+---------+------   840
      TACCTGTACACAGGTAAGCTGTGACAGCGATCTTGAAGACTGCGATGACTTAACAGAGGT

      F  C  A  L  F  T  H  D  E  W  I  Q  Y  D  Y  L  Q  S  L  G    300
      TTCTGTGCTTTGTTCACTCACGACGAATGGATCCAATACGACTACTTGCAAAGCTTGGGT
 841  ---+---------+---------+---------+---------+---------+------   900
      AAGACACGAAACAAGTGAGTGCTGCTTACCTAGGTTATGCTGATGAACGTTTCGAACCCA

      K  Y  Y  G  Y  G  A  G  N  P  L  G  P  A  Q  G  V  G  F  A    320
      AAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTGTTGGTTTCGCT
 901  ---+---------+---------+---------+---------+---------+------   960
      TTCATGATGCCAATGCCACGACCATTGGGTAACCCAGGTCGAGTTCCACAACCAAAGCGA

      N  E  L  I  A  R  L  T  H  S  P  V  Q  D  H  T  S  T  N  H    340
      AACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAAGACCACACTTCTACTAACCAC
 961  ---+---------+---------+---------+---------+---------+------  1020
      TTGCTTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAAGATGATTGGTG

      T  L  D  S  N  P  A  T  F  P  L  N  A  T  L  Y  A  D  F  S    360
      ACTTTGGACTCTAACCCAGCTACTTTCCCATTGAACGCTACTTTGTACGCTGACTTCTCT
1021  ---+---------+---------+---------+---------+---------+------  1080
      TGAAACCTGAGATTGGGTCGATGAAAGGGTAACTTGCGATGAAACATGCGACTGAAGAGA

      H  D  N  T  M  I  S  I  F  F  A  L  G  L  Y  N  G  T  K  P    380
      CACGACAACACTATGATATCTATTTTCTTCGCTTTGGGTTTGTACAACGGTACCAAGCCA
1081  ---+---------+---------+---------+---------+---------+------  1140
      GTGCTGTTGTGATACTATAGATAAAAGAAGCGAAACCCAAACATGTTGCCATGGTTCGGT

      L  S  T  T  S  V  E  S  I  E  E  T  D  G  Y  S  A  S  W  T    400
      TTGTCTACTACTTCTGTTGAATCTATTGAAGAAACTGACGGTTACTCTGCTTCTTGGACT
1141  ---+---------+---------+---------+---------+---------+------  1200
      AACAGATGATGAAGACAACTTAGATAACTTCTTTGACTGCCAATGAGACGAAGAACCTGA

      V  P  F  A  A  R  A  Y  V  E  M  M  Q  C  Q  A  E  K  E  P    420
      GTTCCATTCGCTGCTAGAGCTTACGTTGAAATGATGCAATGTCAAGCTGAAAAGGAACCA
1201  ---+---------+---------+---------+---------+---------+------  1260
      CAAGGTAAGCGACGATCTCGAATGCAACTTTACTACGTTACAGTTCGACTTTTCCTTGGT

      L  V  R  V  L  V  N  D  R  V  V  P  L  H  G  C  A  V  D  K    440
      TTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTGCTGTTGACAAG
1261  ---+---------+---------+---------+---------+---------+------  1320
      AACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACACGACAACTGTTC

      L  G  R  C  K  R  D  D  F  V  E  G  L  S  F  A  R  S  G  G    460
```

Fig. 9B

```
     TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTAGATCTGGTGGT
1321 ---+---------+---------+---------+---------+---------+------ 1380
     AACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGATCTAGACCACCA

     N  W  A  E  C  F  A  *    467
     AACTGGGCTGAATGTTTCGCTTAA
1381 ---+---------+---------+ 1410
     TTGACCCGACTTACAAAGCGAATT
```

Fig. 9C

```
        M  G  V  F  V  V  L  L  S  I  A  T  L  F  G  S  T  S  G  T     20
        ATGGGCGTGTTCGTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCCACATCCGGTACC
      1 ---------+---------+---------+---------+---------+---------+  60
        TACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGTGTAGGCCATGG

        A  L  G  P  R  G  N  S  H  S  C  D  T  V  D  G  G  Y  Q  C     40
        GCCTTGGGTCCTCGTGGTAACTCTCACTCTTGTGACACTGTTGACGGTGGTTACCAATGT
     61 ---------+---------+---------+---------+---------+---------+ 120
        CGGAACCCAGGAGCACCATTGAGAGTGAGAACACTGTGACAACTGCCACCAATGGTTACA

A       F  P  E  I  S  H  L  W  G  T  Y  S  P  F  F  S  L  A  D  E     60
        TTCCCAGAAATTTCTCACTTGTGGGGTACATACTCTCCATTCTTCTCTTTGGCTGACGAA
    121 ---------+---------+---------+---------+---------+---------+ 180
        AAGGGTCTTTAAAGAGTGAACACCCCATGTATGAGAGGTAAGAAGAGAAACCGACTGCTT

        S  A  I  S  P  D  V  P  K  G  C  R  V  T  F  V  Q  V  L  S     80
        TCTGCTATTTCTCCAGACGTTCCAAAGGGTTGTAGAGTTACTTTCGTTCAAGTTTTGTCT
    181 ---------+---------+---------+---------+---------+---------+ 240
        AGACGATAAAGAGGTCTGCAAGGTTTCCCAACATCTCAATGAAAGCAAGTTCAAAACAGA

        R  H  G  A  R  Y  P  T  S  S  A  S  K  A  Y  S  A  L  I  E    100
        AGACACGGTGCTAGATACCCAACTTCTTCTGCGTCTAAGGCGTACTCTGCTTTGATTGAA
    241 ---------+---------+---------+---------+---------+---------+ 300
        TCTGTGCCACGATCTATGGGTTGAAGAAGACGCAGATTCCGCATGAGACGAAACTAACTT

        A  I  Q  K  N  A  T  A  F  K  G  K  Y  A  F  L  K  T  Y  N    120
        GCTATTCAAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGAAGACTTACAAC
    301 ---------+---------+---------+---------+---------+---------+ 360
        CGATAAGTTTTCTTGCGATGACGAAAGTTCCCATTCATGCGAAAGAACTTCTGAATGTTG

A       Y  T  L  G  A  D  D  L  T  P  F  G  E  Q  Q  M  V  N  S  G    140
        TACACTTTGGGTGCTGACGACTTGACTCCATTCGGTGAACAACAAATGGTTAACTCTGGT
    361 ---------+---------+---------+---------+---------+---------+ 420
        ATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTGTTGTTTACCAATTGAGACCA

        I  K  F  Y  R  R  Y  K  A  L  A  R  K  I  V  P  F  I  R  A    160
        ATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCATTCATTAGAGCT
    421 ---------+---------+---------+---------+---------+---------+ 480
        TAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTAAGTAATCTCGA

        S  G  S  D  R  V  I  A  S  A  E  K  F  I  E  G  F  Q  S  A    180
        TCTGGTTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTTTCCAATCTGCT
    481 ---------+---------+---------+---------+---------+---------+ 540
        AGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAAAGGTTAGACGA

        K  L  A  D  P  G  A  N  P  H  Q  A  S  P  V  I  N  V  I  I    200
        AAGTTGGCTGACCCAGGTGCTAACCCACACCAAGCTTCTCCAGTTATTAACGTTATTATT
    541 ---------+---------+---------+---------+---------+---------+ 600
        TTCAACCGACTGGGTCCACGATTGGGTGTGGTTCGAAGAGGTCAATAATTGCAATAATAA

        P  E  G  A  G  Y  N  N  T  L  D  H  G  L  C  T  A  F  E  E    220
        CCAGAAGGTGCTGGTTACAACAACACTTTGGACCACGGTTTGTGTACTGCTTTCGAAGAA
    601 ---------+---------+---------+---------+---------+---------+ 660
        GGTCTTCCACGACCAATGTTGTTGTGAAACCTGGTGCCAAACACATGACGAAAGCTTCTT
```

Fig. 10A

```
       S  E  L  G  D  D  V  E  A  N  F  T  A  V  F  A  P  P  I  R     240
       TCTGAATTGGGTGACGACGTTGAAGCTAACTTCACTGCTGTTTTCGCTCCACCAATTAGA
 661   ---------+---------+---------+---------+---------+---------+   720
       AGACTTAACCCACTGCTGCAACTTCGATTGAAGTGACGACAAAAGCGAGGTGGTTAATCT

       A  R  L  E  A  H  L  P  G  V  N  L  T  D  E  D  V  V  N  L     260
       GCTAGATTGGAAGCTCACTTGCCAGGTGTTAACTTGACTGACGAAGACGTTGTTAACTTG
 721   ---------+---------+---------+---------+---------+---------+   780
       CGATCTAACCTTCGAGTGAACGGTCCACAATTGAACTGACTGCTTCTGCAACAATTGAAC

       M  D  M  C  P  F  D  T  V  A  R  T  S  D  A  T  Q  L  S  P     280
       ATGGACATGTGTCCATTCGACACTGTTGCTAGAACTTCTGACGCTACTCAATTGTCTCCA
 781   ---------+---------+---------+---------+---------+---------+   840
       TACCTGTACACAGGTAAGCTGTGACAACGATCTTGAAGACTGCGATGAGTTAACAGAGGT

       F  C  D  L  F  T  H  D  E  W  I  Q  Y  D  Y  L  Q  S  L  G     300
       TTCTGTGACTTGTTCACTCACGACGAATGGATTCAATACGACTACTTGCAATCTTTGGGT
 841   ---------+---------+---------+---------+---------+---------+   900
       AAGACACTGAACAAGTGAGTGCTGCTTACCTAAGTTATGCTGATGAACGTTAGAAACCCA

       K  Y  Y  G  Y  G  A  G  N  P  L  G  P  A  Q  G  V  G  F  V     320
       AAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTGTTGGTTTCGTT
 901   ---------+---------+---------+---------+---------+---------+   960
       TTCATGATGCCAATGCCACGACCATTGGGTAACCCAGGTCGAGTTCCACAACCAAAGCAA

       N  E  L  I  A  R  L  T  H  S  P  V  Q  D  H  T  S  T  N  H     340
       AACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAAGACCACACTTCTACTAACCAC
 961   ---------+---------+---------+---------+---------+---------+  1020
       TTGCTTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAAGATGATTGGTG

       T  L  D  S  N  P  A  T  F  P  L  N  A  T  L  Y  A  D  F  S     360
       ACTTTGGACTCTAACCCAGCTACTTTCCCATTGAACGCTACTTTGTACGCTGACTTCTCT
1021   ---------+---------+---------+---------+---------+---------+  1080
       TGAAACCTGAGATTGGGTCGATGAAAGGGTAACTTGCGATGAAACATGCGACTGAAGAGA

       H  D  N  T  M  V  S  I  F  F  A  L  G  L  Y  N  G  T  K  P     380
       CACGACAACACTATGGTTTCTATTTTCTTCGCTTTGGGTTTGTACAACGGTACTAAGCCA
1081   ---------+---------+---------+---------+---------+---------+  1140
       GTGCTGTTGTGATACCAAAGATAAAAGAAGCGAAACCCAAACATGTTGCCATGATTCGGT

       L  S  T  T  S  V  E  S  I  E  E  T  D  G  Y  S  A  S  W  T     400
       TTGTCTACTACTTCTGTTGAATCTATTGAAGAAACTGACGGTTACTCTGCTTCTTGGACT
1141   ---------+---------+---------+---------+---------+---------+  1200
       AACAGATGATGAAGACAACTTAGATAACTTCTTTGACTGCCAATGAGACGAAGAACCTGA

       V  P  F  A  A  R  A  Y  V  E  M  M  Q  C  E  A  E  K  E  P     420
       GTTCCATTCGCTGCTAGAGCTTACGTTGAAATGATGCAATGTGAAGCTGAAAAGGAACCA
1201   ---------+---------+---------+---------+---------+---------+  1260
       CAAGGTAAGCGACGATCTCGAATGCAACTTTACTACGTTACACTTCGACTTTTCCTTGGT

       L  V  R  V  L  V  N  D  R  V  V  P  L  H  G  C  G  V  D  K     440
       TTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTGGTGTTGACAAG
1261   ---------+---------+---------+---------+---------+---------+  1320
       AACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACACCACAACTGTTC

       L  G  R  C  K  R  D  D  F  V  E  G  L  S  F  A  R  S  G  G     460
```

Fig. 10B

```
     TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTAGATCTGGTGGT
1321 ---------+---------+---------+---------+---------+---------+ 1380
     AACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGATCTAGACCACCA

     N  W  E  E  C  F  A  *   467
     AACTGGGAAGAATGTTTCGCTTAA
1381 ---------+---------+---- 1404
     TTGACCCTTCTTACAAAGCGAATT
```

Fig. 10C

```
     M  G  V  F  V  V  L  L  S  I  A  T  L  F  G  S  T  S  G  T   20
     ATGGGGGTTTTCGTCGTTCTATTATCTATCGCGACTCTGTTCGGCAGCACATCGGGCACT
  1 ---------+---------+---------+---------+---------+---------+ 60
     TACCCCCAAAAGCAGCAAGATAATAGATAGCGCTGAGACAAGCCGTCGTGTAGCCCGTGA

     A  L  G  P  R  G  N  H  S  K  S  C  D  T  V  D  L  G  Y  Q   40
     GCGCTGGGCCCCCGTGGAAATCACTCCAAGTCCTGCGATACGGTAGACCTAGGGTACCAG
 61 ---------+---------+---------+---------+---------+---------+ 120
     CGCGACCCGGGGGCACCTTTAGTGAGGTTCAGGACGCTATGCCATCTGGATCCCATGGTC

     C  S  P  A  T  S  H  L  W  G  T  Y  S  P  Y  F  S  L  E  D   60
     TGCTCCCCTGCGACTTCTCATCTATGGGGCACGTACTCGCCATaCTTTTCGCTCGAGGAC
121 ---------+---------+---------+---------+---------+---------+ 180
     ACGAGGGGACGCTGAAGAGTAGATACCCCGtgCATGAGCGGTAtGAAAAGCGAGCTCCTG

     E  L  S  V  S  S  K  L  P  K  D  C  R  I  T  L  V  Q  V  L   80
     GAGCTGTCCGTGTCGAGTAAGCTTCCCAAGGATTGCCGGATCACCTTGGTACAGGTGCTA
181 ---------+---------+---------+---------+---------+---------+ 240
     CTCGACAGGCACAGCTCATTCGAAGGGTTCCTAACGGCCTAGTGGAACCATGTCCACGAT

     S  R  H  G  A  R  Y  P  T  S  S  K  S  K  K  Y  K  K  L  I   100
     TCGCGCCATGGAGCGCGGTACCCAACCAGCTCCAAGAGCAAAAAGTATAAGAAGCTTaTt
241 ---------+---------+---------+---------+---------+---------+ 300
     AGCGCGGTACCTCGCGCCATGGGTTGGTCGAGGTTCTCGTTTTTCATATTCTTCGAAtAa

     T  A  I  Q  A  N  A  T  D  F  K  G  K  Y  A  F  L  K  T  Y   120
     ACGGCGATCCAGGCCAATGCCACCGACTTCAAGGGCAAGTacGCCTTTTTGAAGACGTAC
301 ---------+---------+---------+---------+---------+---------+ 360
     TGCCGCTAGGTCCGGTTACGGTGGCTGAAGTTCCCGTTCAtgCGGAAAAACTTCTGCATG

     N  Y  T  L  G  A  D  D  L  T  P  F  G  E  Q  Q  L  V  N  S   140
     AACTATACTCTGGGTGCGGATGACCTCACTCCCTTTGGGGAGCAGCAGCTGGTGAACTCG
361 ---------+---------+---------+---------+---------+---------+ 420
     TTGATATGAGACCCACGCCTACTGGAGTGAGGGAAACCCCTCGTCGTCGACCACTTGAGC

     G  I  K  F  Y  Q  R  Y  K  A  L  A  R  S  V  V  P  F  I  R   160
     GGCATCAAGTTCTACCAGAGGTACAAGGCTCTGGCGCGCAGTGTGGTGCCGTTTATTCGC
421 ---------+---------+---------+---------+---------+---------+ 480
     CCGTAGTTCAAGATGGTCTCCATGTTCCGAGACCGCGCGTCACACCACGGCAAATAAGCG

     A  S  G  S  D  R  V  I  A  S  G  E  K  F  I  E  G  F  Q  Q   180
     GCCTCAGGCTCGGACCGGGTTATTGCTTCGGGAGAGAAGTTCATCGAGGGGTTCCAGCAG
481 ---------+---------+---------+---------+---------+---------+ 540
     CGGAGTCCGAGCCTGGCCCAATAACGAAGCCCTCTCTTCAAGTAGCTCCCCAAGGTCGTC

     A  K  L  A  D  P  G  A  T  N  R  A  A  P  A  I  S  V  I  I   200
     GCGAAGCTGGCTGATCCTGGCGCGACGAACCGCGCCGCTCCGGCGATTAGTGTGATTATT
541 ---------+---------+---------+---------+---------+---------+ 600
     CGCTTCGACCGACTAGGACCGCGCTGCTTGGCGCGGCGAGGCCGCTAATCACACTAATAA

     P  E  S  E  T  F  N  N  T  L  D  H  G  V  C  T  K  F  E  A   220
     CCGGAGAGCGAGACGTTCAACAATACGCTGGACCACGGTGTGTGCACGAAGTTTGAGGCG
601 ---------+---------+---------+---------+---------+---------+ 660
     GGCCTCTCGCTCTGCAAGTTGTTATGCGACCTGGTGCCACACACGTGCTTCAAACTCCGC
```

Fig. 11A

```
      S  Q  L  G  D  E  V  A  A  N  F  T  A  L  F  A  P  D  I  R   240
      AGTCAGCTGGGAGATGAGGTTGCGGCCAATTTCACTGCGCTCTTTGCACCCGACATCCGA
 661  ---------+---------+---------+---------+---------+---------+ 720
      TCAGTCGACCCTCTACTCCAACGCCGGTTAAAGTGACGCGAGAAACGTGGGCTGTAGGCT

      A  R  L  E  K  H  L  P  G  V  T  L  T  D  E  D  V  V  S  L   260
      GCTCGCctCGAGAAGCATCTTCCTGGCGTGACGCTGACAGACGAGGACGTTGTCAGTCTA
 721  ---------+---------+---------+---------+---------+---------+ 780
      CGAGCGgaGCTCTTCGTAGAAGGACCGCACTGCGACTGTCTGCTCCTGCAACAGTCAGAT

      M  D  M  C  P  F  D  T  V  A  R  T  S  D  A  S  Q  L  S  P   280
      ATGGACATGTGTcCGTTTGATACGGTAGCGCGCACCAGCGACGCAAGTCAGCTGTCACCG
 781  ---------+---------+---------+---------+---------+---------+ 840
      TACCTGTACACAgGCAAACTATGCCATCGCGCGTGGTCGCTGCGTTCAGTCGACAGTGGC

      F  C  Q  L  F  T  H  N  E  W  K  K  Y  D  Y  L  Q  S  L  G   300
      TTCTGTCAACTCTTCACTCACAATGAGTGGAAGAAGTACgACTACCTTCAGTCCTTGGGC
 841  ---------+---------+---------+---------+---------+---------+ 900
      AAGACAGTTGAGAAGTGAGTGTTACTCACCTTCTTCATGcTGATGGAAGTCAGGAACCCG

      K  Y  Y  G  Y  G  A  G  N  P  L  G  P  A  Q  G  I  G  F  T   320
      AAGTACTACGGCTACGGCGCAGGCAACCCTCTGGGACCGGCTCAGGGGATAGGGTTCACC
 901  ---------+---------+---------+---------+---------+---------+ 960
      TTCATGATGCCGATGCCGCGTCCGTTGGGAGACCCTGGCCGAGTCCCCTATCCCAAGTGG

      N  E  L  I  A  R  L  T  R  S  P  V  Q  D  H  T  S  T  N  S   340
      AACGAGCTGATTGCCCGGTTGACgCGTTCGCCAGTGCAGGACCACACCAGCACTAACTCG
 961  ---------+---------+---------+---------+---------+---------+ 1020
      TTGCTCGACTAACGGGCCAACTGcGCAAGCGGTCACGTCCTGGTGTGGTCGTGATTGAGC

      T  L  V  S  N  P  A  T  F  P  L  N  A  T  M  Y  V  D  F  S   360
      ACTCTAGTCTCCAACCCGGCCACCTTCCCGTTGAACGCTACCATGTACGTCGACTTTTCA
1021  ---------+---------+---------+---------+---------+---------+ 1080
      TGAGATCAGAGGTTGGGCCGGTGGAAGGGCAACTTGCGATGGTACATGCAGCTGAAAAGT

      H  D  N  S  M  V  S  I  F  F  A  L  G  L  Y  N  G  T  E  P   380
      CACGACAACAGCATGGTTTCCATCTTCTTTGCATTGGGCCTGTACAACGGCACTGAACCC
1081  ---------+---------+---------+---------+---------+---------+ 1140
      GTGCTGTTGTCGTACCAAAGGTAGAAGAAACGTAACCCGGACATGTTGCCGTGACTTGGG

      L  S  R  T  S  V  E  S  A  K  E  L  D  G  Y  S  A  S  W  V   400
      TTGTCCCGGACCTCGGTGGAAAGCGCCAAGGAATTGGATGGGTATTCTGCATCCTGGGTG
1141  ---------+---------+---------+---------+---------+---------+ 1200
      AACAGGGCCTGGAGCCACCTTTCGCGGTTCCTTAACCTACCCATAAGACGTAGGACCCAC

      V  P  F  G  A  R  A  Y  F  E  T  M  Q  C  K  S  E  K  E  P   420
      GTGCCTTTCGGCGCGCGAGCCTACTTCGAGACGATGCAATGCAAGTCGGAAAAGGAGCCT
1201  ---------+---------+---------+---------+---------+---------+ 1260
      CACGGAAAGCCGCGCGCTCGGATGAAGCTCTGCTACGTTACGTTCAGCCTTTTCCTCGGA

      L  V  R  A  L  I  N  D  R  V  V  P  L  H  G  C  D  V  D  K   440
      CTTGTTCGCGCTTTGATTAATGACCGGGTTGTGCCACTGCATGGCTGCGATGTGGACAAG
1261  ---------+---------+---------+---------+---------+---------+ 1320
      GAACAAGCGCGAAACTAATTACTGGCCCAACACGGTGACGTACCGACGCTACACCTGTTC

      L  G  R  C  K  L  N  D  F  V  K  G  L  S  W  A  R  S  G  G   460
```

Fig. 11B

```
     CTGGGGCGATGCAAGCTGAATGACTTTGTCAAGGGATTGAGTTGGGCCAGATCTGGGGGC
1321 ---------+---------+---------+---------+---------+---------+ 1380
     GACCCCGCTACGTTCGACTTACTGAAACAGTTCCCTAACTCAACCCGGTCTAGACCCCCG

     N  W  G  E  C  F  S  *   467
     AACTGGGGAGAGTGCTTTAGTTGA
1381 ---------+---------+---- 1404
     TTGACCCCTCTCACGAAATCAACT
```

Fig. 11C

```
    CP-1
      Eco RI   M  G  V  F  V  V  L  L  S  I  A  T  L  F  G  S  T
    TATATGAATTCATGGGCGTGTTCGTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCCA
  1 ---------+---------+---------+---------+---------+---------+ 60
    ATATACTTAAGTACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGT

      S  G  T  A  L  G  P  R  G  N  S  H  S  C  D  T  V  D  G  G
    CATCCGGTACCGCCTTGGGTCCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTG
 61 ---------+---------+---------+---------+---------+---------+ 120
    GTAGGCCATGGCGGAACCCAGGAGCACCATTAAGAGTGAGAACACTGTGACAACTGCCAC
         CP-2
              CP-3
      Y  Q  C  F  P  E  I  S  H  L  W  G  Q  Y  S  P  Y  F  S  L
    GTTACCAATGTTTCCCAGAAATTTCTCACTTGTGGGGTCAATACTCTCCATACTTCTCTT
121 ---------+---------+---------+---------+---------+---------+ 180
    CAATGGTTACAAAGGGTCTTTAAAGAGTGAACACCCCAGTTATGAGAGGTATGAAGAGAA

      E  D  E  S  A  I  S  P  D  V  P  D  D  C  R  V  T  F  V  Q
    TGGAAGACGAATCTGCTATTTCTCCAGACGTTCCAGACGACTGTAGAGTTACTTTCGTTC
181 ---------+---------+---------+---------+---------+---------+ 240
    ACCTTCTGCTTAGACGATAAAGAGGTCTGCAAGGTCTGCTGACATCTCAATGAAAGCAAG
                   CP-4.7
                        CP-5.7
      V  L  S  R  H  G  A  R  Y  P  T  D  S  K  G  K  K  Y  S  A
    AAGTTTTGTCTAGACACGGTGCTAGATACCCAACTgacTCTAAGggtAAGaagTACTCTG
241 ---------+---------+---------+---------+---------+---------+ 300
    TTCAAAACAGATCTGTGCCACGATCTATGGGTTGActgAGATTCccaTTCttcATGAGAC

      L  I  E  A  I  Q  K  N  A  T  A  F  K  G  K  Y  A  F  L  K
    CTTTGATTGAAGCTATTCAAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGA
301 ---------+---------+---------+---------+---------+---------+ 360
    GAAACTAACTTCGATAAGTTTTCTTGCGATGACGAAAGTTCCCATTCATGCGAAAGAACT
                             CP-6
                                  CP-7
      T  Y  N  Y  T  L  G  A  D  D  L  T  P  F  G  E  N  Q  M  V
    AGACTTACAACTACACTTTGGGTGCTGACGACTTGACTCCATTCGGTGAAAACCAAATGG
361 ---------+---------+---------+---------+---------+---------+ 420
    TCTGAATGTTGATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACC

      N  S  G  I  K  F  Y  R  R  Y  K  A  L  A  R  K  I  V  P  F
    TTAACTCTGGTATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCAT
421 ---------+---------+---------+---------+---------+---------+ 480
    AATTGAGACCATAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTA
                                      CP-8.7
                                           CP-9
      I  R  A  S  G  S  S  R  V  I  A  S  A  E  K  F  I  E  G  F
    TCATTAGAGCTTCTGGTTCTTctAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTT
481 ---------+---------+---------+---------+---------+---------+ 540
    AGTAATCTCGAAGACCAAGAagaTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAA

      Q  S  A  K  L  A  D  P  G  S  Q  P  H  Q  A  S  P  V  I  D
    TCCAATCTGCTAAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTG
541 ---------+---------+---------+---------+---------+---------+ 600
    AGGTTAGACGATTCAACCGACTGGGTCCAAGAGTTGGTGTGGTTCGAAGAGGTCAATAAC
```

Fig. 12A

```
                                                     CP-10.7
                                                          CP-11.7
       V  I  I  S  E  A  S  S  Y  N  N  T  L  D  P  G  T  C  T  A
       ACGTTATTATTtctGAcgctTCTtctTACAACAACACTTTGGACccaGGTACTTGTACTG
   601 ---------+---------+---------+---------+---------+---------+ 660
       TGCAATAATAAagaCTgcgaAGGagaATGTTGTTGTGAAACCTGggtCCATGAACATGAC
```

Fig. 12B

```
      F  E  D  S  E  L  A  D  T  V  E  A  N  F  T  A  L  F  A  P
     CTTTCGAAGACTCTGAATTGgctGACactGTTGAAGCTAACTTCACTGCTTTGTTCGCTC
661  ---------+---------+---------+---------+---------+---------+ 720
     GAAAGCTTCTGAGACTTAACcgaCTGtgaCAACTTCGATTGAAGTGACGAAACAAGCGAG
                                                         CP-12.7

      A  I  R  A  R  L  E  A  D  L  P  G  V  T  L  T  D  T  E  V
     CAGCTATTAGAGCTAGATTGGAAGCTGACTTGCCAGGTGTTACTTTGACTGACactgaaG
721  ---------+---------+---------+---------+---------+---------+ 780
     GTCGATAATCTCGATCTAACCTTCGACTGAACGGTCCACAATGAAACTGACTGtgacttC

     CP-13.7
      T  Y  L  M  D  M  C  S  F  E  T  V  A  R  T  S  D  A  T  E
     TTactTACTTGATGGACATGTGTtctTTCGAAACTGTTGCTAGAACTTCTGACGCTACTG
781  ---------+---------+---------+---------+---------+---------+ 840
     AAtgaATGAACTACCTGTACACAagaAAGCTTTGACAACGATCTTGAAGACTGCGATGAC

      L  S  P  F  C  A  L  F  T  H  D  E  W  R  H  Y  D  Y  L  Q
     AATTGTCTCCATTCTGTGCTTTGTTCACTCACGACGAATGGAGAcacTACGACTACTTGC
841  ---------+---------+---------+---------+---------+---------+ 900
     TTAACAGAGGTAAGACACGAAACAAGTGAGTGCTGCTTACCTCTgtgATGCTGATGAACG
           CP-14.7
                CP-15.7
      S  L  K  K  Y  Y  G  H  G  A  G  N  P  L  G  P  T  Q  G  V
     AATCTTTGaagAAGTACTACGGTcacGGTGCTGGTAACCCATTGGGTCCAactCAAGGTG
901  ---------+---------+---------+---------+---------+---------+ 960
     TTAGAAACttcTTCATGATGCCAgtgCCACGACCATTGGGTAACCCAGGTtgaGTTCCAC

      G  F  A  N  E  L  I  A  R  L  T  R  S  P  V  Q  D  H  T  S
     TTGGTTTCGCTAACGAATTGATTGCTAGATTGACTAGATCTCCAGTTCAAGACCACACTT
961  ---------+---------+---------+---------+---------+---------+ 1020
     AACCAAAGCGATTGCTTAACTAACGATCTAACTGATCTAGAGGTCAAGTTCTGGTGTGAA
                     CP-16
                        CP-17.7
      T  N  H  T  L  D  S  N  P  A  T  F  P  L  N  A  T  L  Y  A
     CTACTAACCACACTTTGGACTCTAACCCAGCTACTTTCCCATTGAACGCTACTTTGTACG
1021 ---------+---------+---------+---------+---------+---------+ 1080
     GATGATTGGTGTGAAACCTGAGATTGGGTCGATGAAAGGGTAACTTGCGATGAAACATGC

      D  F  S  H  D  N  G  I  I  S  I  F  F  A  L  G  L  Y  N  G
     CTGACTTCTCTCACGACAACggtattATTTCTATTTTCTTCGCTTTGGGTTTGTACAACG
1081 ---------+---------+---------+---------+---------+---------+ 1140
     GACTGAAGAGAGTGCTGTTGccataaTAAAGATAAAAGAAGCGAAACCCAAACATGTTGC
                              CP-18.7
                                   CP-19.7
      T  A  P  L  S  T  T  S  V  E  S  I  E  E  T  D  G  Y  S  S
     GTACTGCTCCATTGTCTACTACTTCTGTTGAATCTATTGAAGAAACTGACGGTTACTCTt
1141 ---------+---------+---------+---------+---------+---------+ 1200
     CATGACGAGGTAACAGATGATGAAGACAACTTAGATAACTTCTTTGACTGCCAATGAGAa

      A  W  T  V  P  F  A  S  R  A  Y  V  E  M  M  Q  C  Q  A  E
     ctgctTGGACTGTTCCATTCgcttctAGAGCTTACGTTGAAATGATGCAATGTCAAGCTG
1201 ---------+---------+---------+---------+---------+---------+ 1260
     gacgaACCTGACAAGGTAAGcgaagaTCTCGAATGCAACTTTACTACGTTACAGTTCGAC
                                        CP-20
                                             CP-21
```

Fig. 12C

```
        K  E  P  L  V  R  V  L  V  N  D  R  V  V  P  L  H  G  C  A
     AAAAGGAACCATTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTG
1261 ---------+---------+---------+---------+---------+---------+ 1320
     TTTTCCTTGGTAACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACAC
```

Fig. 12D

```
        V  D  K  L  G  R  C  K  R  D  D  F  V  E  G  L  S  F  A  R
     CTGTTGACAAGTTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTA
1321 ---------+---------+---------+---------+---------+---------+ 1380
     GACAACTGTTCAACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGAT
                                                  CP-22

        S  G  G  N  W  A  E  C  F  A  *  Eco RI
     GATCTGGTGGTAACTGGGCTGAATGTTTCGCTTAAGAATTCATATA
1381 ---------+---------+---------+---------+------ 1426
     CTAGACCACCATTGACCCGACTTACAAAGCGAATTCTTAAGTATAT
```

Fig. 12E