Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2319658 C 2010/06/15

(11)(21) 2 319 658

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(22) Date de dépôt/Filing Date: 2000/10/06

(41) Mise à la disp. pub./Open to Public Insp.: 2001/04/11

(45) Date de délivrance/Issue Date: 2010/06/15

(30) Priorités/Priorities: 1999/10/11 (EP99120289.6);
2000/09/08 (EP00119676.5)

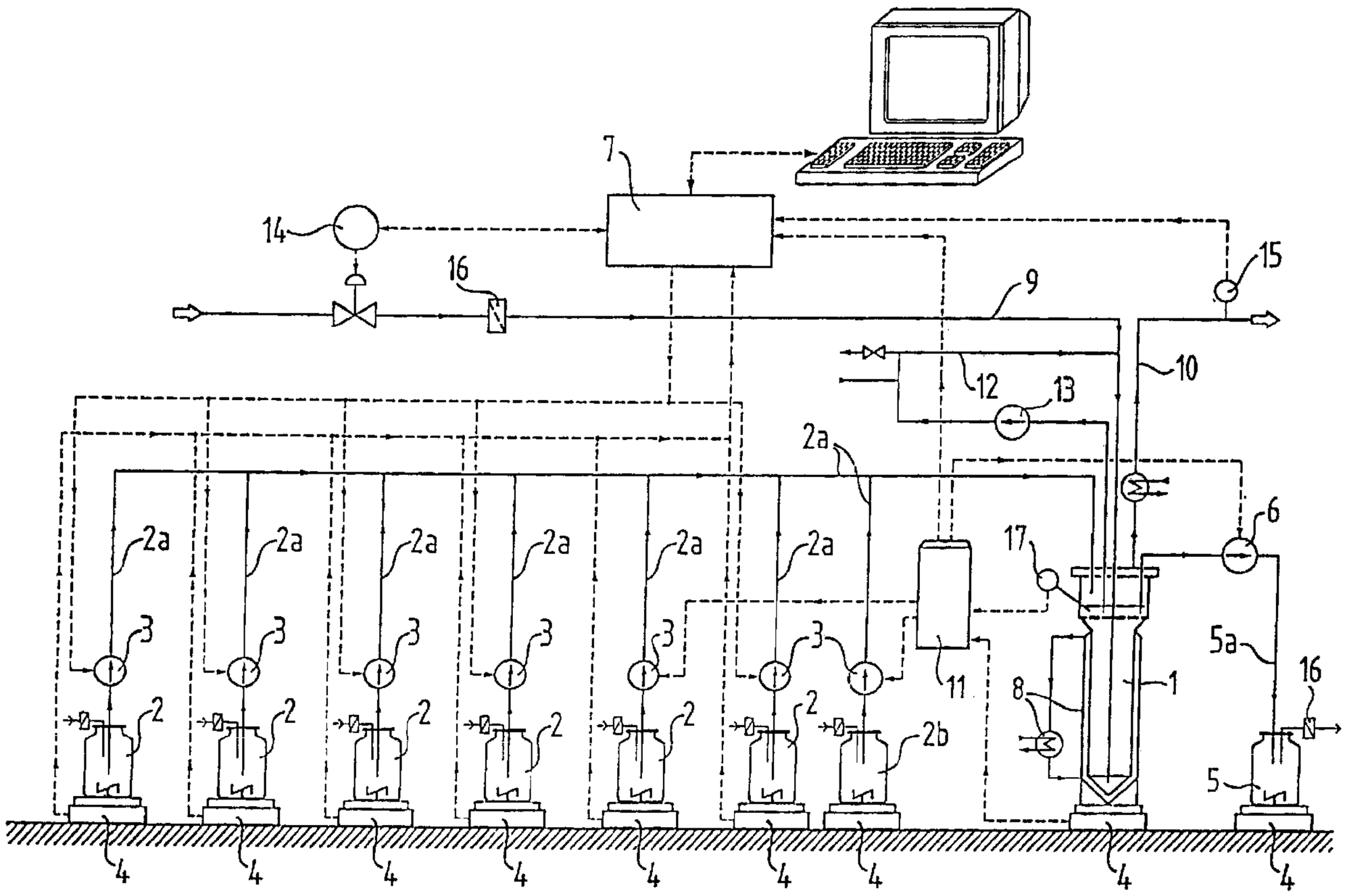
(51) Cl.Int./Int.Cl. *C12M 1/00* (2006.01),
C12M 1/36 (2006.01), *C12P 21/00* (2006.01)

(72) Inventeurs/Inventors:
BARTOK, ATTILA, CH;
MUEH, THORSTEN, DE;
RUECKEL, MARKUS, DE

(73) Propriétaire/Owner:
DSM IP ASSETS B.V., NL

(74) Agent: SMART & BIGGAR

(54) Titre : PROCESSUS DE FERMENTATION CONTINU
(54) Title: CONTINUOUS FERMENTATION PROCESS



(57) Abrégé/Abstract:
The invention is concerned with a continuous process for the manufacture of proteins by means of protein-producing microorganism in which process the microorganism is optionally immobilized on a solid carrier and/or the nutrients and other agents required for the growth of the microorganism and the optimal production of protein are fed into the reactor individually at a constant dilution rate. Furthermore, the invention is concerned with a process for the manufacture of proteins using a fermentation assembly.

Abstract

The invention is concerned with a continuous process for the manufacture of proteins by means of protein-producing microorganism in which process the microorganism is optionally immobilized on a solid carrier and/or the nutrients and other agents required for the growth of the microorganism and the optimal production of protein are fed into the reactor individually at a constant dilution rate. Furthermore, the invention is concerned with a process for the manufacture of proteins using a fermentation assembly.

Continuous Fermentation Process

Case 20486

The present invention relates to a continuous process for the manufacture of proteins.

In accordance with the present invention it has been found that splitting of cultivation media used in a continuous fermentation process allows to study the influence on growth and metabolite-production of microorganisms and thus to determine optimal conditions for the fermentation process. A continuously delivered fermentation medium can generally be split into as many fractions as it contains ingredients. Examples of such ingredients are carbon, nitrogen, phosphorus and sulfur sources as well as vitamins and complex substrates such as corn steep, yeast extract and other natural products. Furthermore, every required mineral, micro- or trace element can be provided separately as a solution of a water-soluble salt, such as a chloride, sulfate or nitrate. In this manner a fermentation medium of any desired composition can be obtained, provided that the desired amounts of the ingredients are (water)-soluble and no disturbing interactions (e.g., precipitation, reaction) occur in the individual feed solutions or in the fermentation medium.

In one aspect, the present invention is concerned with a continuous process for the manufacture of proteins by means of protein-producing microorganism.

SD/vs/31.8.2000

More particularly, the invention is concerned with a continuous process for the manufacture of proteins by means of protein-producing microorganism in which process the microorganism is optionally immobilized on a solid carrier and/or the nutrients and other agents required for the growth of the microorganism and the optimal production of protein are fed into the reactor individually at a constant dilution rate.

In a preferred aspect, the invention is concerned with a process for the manufacture of proteins using a fermentation assembly that comprises

a vessel suitable for carrying out reactions involving living or inactivated cells;

at least two storage flasks connected to said vessel for supply of liquids and means to transport said liquids from said storage flasks to said vessel;

individual appliances monitoring the supply of the contents of said storage flasks to said vessel;

a harvest flask connected to said vessel and means to transport fermentation broth from said vessel to said harvest flask; and

a device for controlling and maintaining a constant dilution rate in said vessel with varying rates of individual supply of liquid from said storage flasks to said vessel.

Any conventional fermentation vessel can be used for the purpose of this invention. The vessel may be made of materials such as stainless steel, glass or ceramics and may have a volume of from e.g., l00 ml to 2500 m^3 although these figures are not critical to the invention. For continuous operation the inside of the vessel is optionally equipped with, e.g., a receptacle or sieve plate for uptake of immobilized cells. Further, the fermentation vessel is connected to a series of storage flasks that contain nutrient solutions and solutions

for maintaining and controlling a desired pH and other parameters, such as foam formation, redox potential etc. in the fermentation broth. Depending on the particular needs of the fermentation, there may be separate storage flasks for individual supply of substrates that serve as carbon or nitrogen or mineral source for the living cells.

It has been found in accordance with the invention that the process is advantageously carried out at a constant dilution rate in the fermentation vessel. As used herein, the term "dilution rate" denotes the total volume of liquids supplied to the fermentation vessel per volume of the fermentation vessel per hour [h^{-1}].

Accordingly, it is a particular feature of the present invention to carry out the fermentation process at a constant dilution rate in the fermentation vessel while varying the supply of individual nutrient components or other additives during the fermentation process. To facilitate this task a storage flask containing an inert component, e.g., water is optionally provided that allows to complement the supply of liquids thus keeping the total supply of liquid constant.

The assembly that is preferably used to carry out the process of this invention further comprises means to transport the individual components of the fermentation medium from the storage flasks to the fermentation vessel, and appliances for monitoring the amount of liquid supplied to the fermentation vessel. Every combination of measuring instruments (e.g., volumetric or mass flow rate by either gravimetric, anemometric, magnetic, ultrasonic, Venturi, J, cross-relation, thermal, Coriolis, radiometric) and transfer units (e.g., pumps or pressure difference) can be used for this purpose. Additionally, every transfer unit can be applied as a dosing unit (e.g., gear, peristaltic, piston, membrane or excenter pump). For operation on small scale the supply is suitably monitored by weighing the storage flasks that contain nutrient or additive solutions in a predetermined concentration.

The device for controlling and maintaining a constant dilution rate in the fermentation vessel is suitably a system comprising a measuring instrument that monitors the flow from the storage flasks and a controlling unit, e.g., a computer-software control that calculates the actual mass flow rates, compares them to the desired value and adjusts the pump

setting accordingly. An appropriate system is, e.g., the Process Automation System, National Instruments, Bridge View, USA, for Windows NT 4.0 (represented by National Instruments, Sonnenbergstrasse 53, 5408 Ennetbaden, Switzerland) that is connected to the various operating units (scales, pumps) through a serial-interface box (Rocket Port, Comtrol Europe Ltd, Great Britain, represented by Technosoftware AG Rothackerstrasse 13, 5702 Niederlenz, Switzerland).

An assembly that can be used in the process of this invention is depicted in Figure 1.

The fermentation vessel 1 (Fermentor) is equipped with inlet tubes 2a from storage flasks 2 (suitably equipped with a stirrer) for supply of salt solution (Salts), nutrient solution (Nutrients), particular substrates (Substrate 1 and Substrate 2) for supply of, e.g., distinct carbon sources, agent for controlling the pH (Base), water for controlling a constant dilution rate, and antifoam. Pumps 3 transport liquids from the storage flasks 2 to the fermentor 1. Scales 4 monitor the amount of liquids supplied to and discharged from the fermentor. Further, the fermentor has inlet tubes 9 for oxygen supply and outlet tubes 10 for exhaust controlled by untits 14 and 15. Pump 6 discharges fermentation broth via outlet tubes 5a to a harvest flask 5. A main controlling unit 7 monitors and steers the overall process. Controlling unit 11 monitors and steers individual control systems 17 for temperature, pH, gas pressure, fermentor content and supply of antifoam agents. Circuit 12 including pump 13 is used for taking samples from the fermentation broth and for providing a controlled gas flow for moving the fermentation broth. Inlet and outlet gas flow is controlled by flow control 14 and 15. Sterile filters 16 are provided optionally. Optionally, the fermentation vessel 1 is equipped with a thermostating unit 8.

In the process of the present invention, any protein-producing microorganism either natural, e.g. fungal origin or bacterial origin or microorganisms which have been transformed by protein encoding DNA whereby such transformed microorganisms can be bacteria or fungi or yeasts, preferably from the genus Peniophora, Aspergillus, Hansenula or Pichia, especially *Aspergillus niger*, *Aspergillus awanari*, *Aspergillus sojae*, *Aspergillus oryzae* or *Hansenula polymorpha* or *Pichia pastoris*.

In this context, the skilled person in the art selects such a protein-producing microorganism which is known to be useful for the production of a desired protein.

In a preferred embodiment of the present invention the protein is selected from the group consisting of proteins having the activity of an enzyme such as catalase, lactase, phenoloxidase, oxidase, oxidoreductase, glucanase cellulase, xylanase and other polysaccharide, peroxidase, lipase, hydrolase, esterase, cutinase, protease and other proteolytic enzymes, aminopeptidase, carboxypeptidase, phytase, lyase, pectinase and other pectinolytic enzymes, amylase, glucosidase, mannosidase, isomerase, invertase, transferase, ribonuclease, chitinase, and desoxyribonuclease. Furthermore, in a preferred embodiment of the present invention the protein is selected from the group of therapeutic proteins such as antibodies, vaccines, antigens, or of antibacterial and/or health-beneficial proteins such as lactoternin, lactoperoxidase or lysozyme.

It will be understood by those skilled in the art that the term "activity" includes not only native activities referring to naturally occuring enzymes or therapeutic functions, but also those activities or functions which have been modified by amino acid substitutions, deletions, additions, or other modifications which may be made to enhance or modify the desired activity, or the thermostability, pH tolerance and/or further properties.

In a most preferred embodiment of the invention the selected protein is a protein having the activity of a phytase.

Examples of proteins having the activity of a phytase are described in EP 684 313, EP 897 010, EP 897 985 or in Examples 6 to 16 and Figures 2 – 22 of the present invention.

Figure 2: Design of the consensus phytase sequence. The letters represent the amino acid residues in the one-letter code. The following sequences were used for the alignment: *phyA* from *Aspergillus terreus* 9A-1 (Mitchell et al, 1997; from amino acid (aa) 27), *phyA* from *A. terreus* cbs116.46; (van Loon et al., 1998; from aa 27), *phyA* from *Aspergillus niger* var. *awamori* (Piddington et al, 1993; from aa 27), *phyA* from *A. niger* T213; Mitchell et al. 1997 from aa 27), *phyA* from *A. niger* strain NRRL3135 (van Hartingsveldt et al, 1993; from aa 27), *phyA* from *Aspergillus fumigatus* ATCC 13073 (Pasamontes et al, 1997; from aa 25), *phyA* from *A. fumigatus* ATCC 32722 (EP 897 985; Figur 1; from aa 27), *phyA* from *A.*

fumigatus ATCC 58128 (EP 897 985; Figur 1; from aa 27), *phyA* from *A. fumigatus* ATCC 26906 (EP 897 985; Figur 1; from aa 27), *phyA* from *A. fumigatus* ATCC 32239 (EP 897 985; Figur 1; from aa 30), *phy*A from *Emericella nidulans* (Pasamontes et al, 1997a; from aa 25), *phy*A from *Talaromyces thermophilus* (Pasamontes et al, 1997a; from aa 24), and *phy*A from *Myceliophthora thermophila* (Mitchell et al, 1997; from aa 19). The alignment was calculated using the program PILEUP. The location of the gaps was refined by hand. Capitalized amino acid residues in the alignment at a given position belong to the amino acid coalition that establish the consensus residue. In bold, beneath the calculated consensus sequence, the amino acid sequence of the finally constructed consensus phytase (Fcp) is shown. The gaps in the calculated consensus sequence were filled by hand according to principals stated in Example 6.

Figure 3: DNA sequence of the consensus phytase-1 gene (*fcp*) and of the primers used for the gene construction. The calculated amino acid sequence (Figure 2) was converted into a DNA sequence using the program BACKTRANSLATE (Devereux *et al.*, 1984) and the codon frequency table of highly expressed yeast genes (GCG program package, 9.0). The signal peptide of the phytase from *A. terreus* cbs.116.46 was fused to the *N*-terminus. The bold bases represent the sequences of the oligonucleotides used to generate the gene. The names of the respective oligonucleotides are alternately noted above or below the sequence. The underlined bases represent the start and stop codon of the gene. The bases written in italics show the two introduced *Eco* RI sites.

Figure 4: Alignment and consensus sequence of five *Basidiomycetes* phytases. The letters represent the amino acid residues in the one-letter code. The amino acid sequences of the phytases from *Paxillus involutus*, phyA1 (aa 21) and phyA2 (aa 21, WO 98/28409), *Trametes pubescens* (aa 24, WO 98/28409), *Agrocybe pediades* (aa 19, WO 98/28409), and *Peniophora lycii* (aa 21, WO 98/28409) starting with the amino acid residues mentioned in parentheses, were used for the alignment and the calculation of the corresponding consensus sequence called "Basidio" (Example 7). The alignment was performed by the program PILEPUP. The location of the gaps was refined by hand. The consensus sequence was calculated by the program PRETTY. While a vote weight of 0.5 was assigned to the two *P. involutus* phytases, all other genes were used with a vote weight of 1.0 for the consensus sequence calculation. At positions, where the program was not able to determine a consensus residue, the Basidio sequence contains a dash. Capitalized amino acid residues in the alignment at a given position belong to the amino acid coalition that establish the consensus residue.

Figure 5: Design of consensus phytase-10 amino acid sequence. Adding the phytase sequence of *Thermomyces lanuginosus* (Berka *et al.*, 1998) and the consensus sequence of the phytases from five *Basidiomycetes* to the alignment of Figure 2, an improved consensus sequence was calculated by the program PRETTY. Additionally, the amino acid sequence of *A. niger* T213 was omitted; therefore, a vote weight of 0.5 was used for the remaining *A. niger* phytase sequences. For further information see Example 8.

Figure 6: DNA and amino acid sequence of consensus phytase-10. The amino acid sequence is written above the corresponding DNA sequences using the one-letter code. The sequence of the oligonucleotides which were used to assemble the gene are in bold letters. The labels of oligonucleotides and the amino acids which were changed compared to those for consensus phytase -1 are underlined. The *fcp*10 gene was assembled from the following oligonucleotides: CP-1, CP-2, CP-3.10, CP-4.10, CP-5.10, CP-6, CP-7.10, CP-8.10, CP-9.10, CP-10.10, CP-11.10, CP-12.10, CP-13.10, CP-14.10, CP-15.10, CP-16.10, CP-17.10, CP18.10, CP-19.10, CP-20.10, CP-21.10, CP-22.10. The newly synthesized oligonucleotides are additionally marked by number 10. The phytase contains the following 32 exchanges relative to consensus phytase -1: Y54F, **E58A**, D69K, D70G, A94K, N134Q, I158V, S187A, Q188N, **D197N**, S204A, T214L, D220E, L234V, A238P, D246H, T251N, Y259N, **E267D**, E277Q, A283D, **R291I**, A320V, **R329H**, **S364T**, I366V, A379K, S396A, **G404A**, Q415E, A437G, A463E. The mutations accentuated in bold letters revealed a stabilizing effect on consensus phytase-1 when tested as single mutations in consensus phytase-1.

Figure 7: Alignment for the design of consensus phytase-11. In contrast to the design of consensus phytase-10, for the design of the amino acid sequence of consensus phytase-11, all *Basidiomycete* phytases were used as independent sequences using an assigned vote weight of 0.2 for each *Basidiomycete* sequence. Additionally, the amino acid sequence of *A. niger* T213 phytase was used in that alignment, again.

Figure 8: DNA and amino acid sequence of consensus phytase-1-thermo[8]-Q50T-K91A. The amino acid sequence is written above the corresponding DNA sequence using the one-letter code. The replaced amino acid residues are underlined. The stop codon of the gene is marked by a star (*).

Figure 9: DNA and amino acid sequence of consensus phytase-10-thermo[3]-Q50T-K91A. The amino acid sequence is written above the corresponding DNA sequence using the

one-letter code. The replaced amino acid residues are underlined. The stop codon of the gene is marked by a star (*).

Figure 10: DNA and amino acid sequence of *A. fumigatus* ATCC 13073 phytase α-mutant. The amino acid sequence is written above the corresponding DNA sequence using the one-letter code. The replaced amino acid residues are underlined. The stop codon of the gene is marked by a star (*).

Figure 11: DNA and amino acid sequence of consensus phytase-7. The amino acids are written above the corresponding DNA sequence using the one-letter code. The sequences of the oligonucleotides used to assemble the gene are in bold letters. Oligonucleotides and amino acids that were exchanged are underlined and their corresponding triplets are highlighted in small cases. The *fcp7* gene was assembled from the following oligonucleotides: CP-1, CP-2, CP-3, CP-4.7, CP-5.7, CP-6, CP-7, CP-8.7, CP-9, CP-10.7, CP-11.7, CP-12.7, CP-13.7, CP-14.7, CP-15.7, CP-16, CP-17.7, CP-18.7, CP-19.7, CP-20, CP-21, CP-22. The newly synthesized oligonucleotides are additionally marked by number 7. The phytase contains the following 24 exchanges in comparison to the original 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.

Figure 12: Differential scanning calorimetry (DSC) of consensus phytase-1 and consensus phytase-10. The protein samples were concentrated to ca. 50-60 mg/ml and extensively dialyzed against 10 mM sodium acetate, pH 5.0. A constant heating rate of 10 °C/min was applied up to 95 °C. DSC of consensus phytase-10 (upper graph) yielded a melting temperature of 85.4 °C, which is 7.3 °C higher than the melting point of consensus phytase-1 (78.1 °C, lower graph).

Figure 13: Differential scanning calorimetry (DSC) of consensus phytase-10-thermo[3]-Q50T and consensus phytase-10-thermo[3]-Q50T-K91A. The protein samples were concentrated to ca. 50-60 mg/ml and extensively dialyzed against 10 mM sodium acetate, pH 5.0. A constant heating rate of 10 °C/min was applied up to 95 °C. DSC of consensus phytase-10-thermo-[3]-Q50T (upper graph) yielded a melting temperature of 88.6 °C, while the melting point of consensus phytase-10-thermo[3]-Q50T-K91A was found at 89.3 °C.

Figure 14: Comparison of the temperature optimum between consensus phytase-1, consensus phytase-10 and consensus phytase-10-thermo[3]-Q50T. For the determination of the temperature optimum, the phytase standard assay was performed at a series of temperatures between 37 and 86 °C. The diluted supernatant of transformed *S. cerevisiae* strains was used for the determination. The other components of the supernatant showed no influence on the determination of the temperature optimum: Δ, consensus phytase-1; ◊, consensus phytase-10; ■, consensus phytase 10-thermo[3]-Q50T.

Figure 15: pH-dependent activity profile and substrate specificity of consensus phytase-10 and its variants thermo[3]-Q50T and thermo[3]-Q50T-K91A. Graph a) shows the pH-dependent activity profile of consensus phytase-10 (□), consensus phytase-10-thermo[3]-Q50T (Δ), and consensus phytase-10-thermo[3]-Q50T-K91A (Δ).The phytase activity was determined using the standard assay in appropriate buffers (see Example 15) at different pH-values. Graph b) shows the corresponding substrate specificity tested by replacement of phytate by the indicated compounds in the standard assay; open bars, consensus phytase-10 (white bars, consensus phytase-10-thermo-Q50T; dark bars, consensus phytase-10-thermo-Q50T-K91A). The numbers correspond to the following compounds: 1, phytate; 2, *p*-nitrophenyl phosphate; 3, phenyl phosphate; 4, fructose-1,6-bisphosphate; 5, fructose-6-phosphate; 6, glucose-6-phosphate; 7, ribose-5-phosphate; 8, DL-glycerol-3-phosphate; 9, glycerol-2-phosphate; 10, 3-phosphoglycerate; 11, phosphoenolpyruvate; 12, AMP; 13, ADP; 14, ATP.

Figure 16: pH-dependent activity profile and substrate specificity of consensus phytase-1-thermo[8]-Q50T and of consensus phytase-1-thermo[8]-Q50T-K91A.. Graph a) shows the pH-dependent activity profile of the Q50T- (■) and the Q50T-K91A-variant (Δ).The phytase activity was determined using the standard assay in appropriate buffers (see Example 15) at different pH-values Graph b) shows the corresponding substrate specificities tested by replacement of phytate by the indicated compounds in the standard assay (open bars, consensus phytase-1-thermo[8]-Q50T; filled bars, consensus phytase-1-thermo[8]-Q50T-K91A). The substrates are listed in the legend of Figure 15.

Figure 17: Differential scanning calorimetry (DSC) of consensus phytase-1-thermo[8]-Q50T and consensus phytase-1-thermo[8]-Q50T-K91A. The protein samples were concentrated to ca. 50-60 mg/ml and extensively dialyzed against 10 mM sodium acetate, pH 5.0. A constant heating rate of 10 °C/min was applied up to 95 °C. DSC of consensus phytase-1-thermo[8]-Q50T (upper graph) showed a melting temperature of 84.7 °C, while the melting point of consensus phytase-1-thermo[8]-Q50T-K91A was found at 85.7 °C.

Figure 18: Comparison of the temperature optimum between consensus phytase-1, consensus phytase-1-thermo[3] and consensus phytase-1-thermo[8]. For the determination of the temperature optimum, the phytase standard assay was performed at a series of temperatures between 37 and 86 °C. Purified protein from the supernatant of transformed *S. cerevisiae* strains was used for the determination. O, consensus phytase-1;□, consensus phytase-1-thermo[3]; Δ, consensus phytase -1-thermo[8].

Figure 19: Comparison of the pH-dependent activity profile and substrate specificity of consensus phytase-1, consensus phytase-7, and of the phytase from *A. niger* NRRL 3135.. Graph a) shows the pH-dependent activity profile of consensus phytase-1 (■), the phytase from *A. niger* NRRL 3135 (O), and of consensus phytase-7 (Δ).The phytase activity was determined using the standard assay in appropriate buffers (see Example 15) at different pH-values Graph b) shows the corresponding substrate specificity tested by replacement of phytate by the indicated compounds in the standard assay (black bars, *A. niger* NRRL 3135 phytase; grey bars, consensus phytase-1, dashed bars, consensus phytase-7). The substrates are listed in the legend of Figure 15.

Figure 20: Differential scanning calorimetry (DSC) of the phytase from A. fumigatus ATCC 13073 and of its stabilized α-mutant, which contains the following amino acid exchanges: F55Y, V100I, F114Y, A243L, S265P, N294D.

The protein samples were concentrated to ca. 50-60 mg/ml and extensively dialyzed against 10 mM sodium acetate, pH 5.0. A constant heating rate of 10 °C/min was applied up to 95 °C. DSC of consensus *A. fumigatus* 13073 phytase (lower graph) revealed a melting temperature of 62.5 °C, while the melting point of the α-mutant was found at 67.0 °C

Figure 21: Comparison of the temperature optimum of *A. fumigatus* 13073 wild-type phytase, its α-mutant, and a further stabilized α-mutant (E59A-S126N-R329H-S364T-G404A). For the determination of the temperature optimum, the phytase standard assay was performed at a series of temperatures between 37 and 75 °C. The diluted supernatants of transformed *S. cerevisiae* strains were used for the determination. The other components of the supernatant showed no influence on the determination of the temperature optimum. O, *A. fumigatus* ATCC 13073 phytase; Δ, *A. fumigatus* ATCC 13073 α-mutant; □, *A. fumigatus* ATCC 13073 alpha-mutant-(E59A-S126N-R329H-S364T-

G404A)-Q27T; ■ , *A. fumigatus* ATCC 13073 α-mutant-(E59A-S126N-R329H-S364T-G404A)-Q27T-K68A. The mutations Q51T and K92A in the *A. furnigatus* α-mutants correspond to -1 Q50T and K91A in consensus phytase, respectively.

Figure 22: Amino acid sequence of consensus phytase -12 (consphy12) which contains a number of active site residues transferred from the "basidio" consensus sequence to consensus phytase-10-thermo[3]-Q50T-K91A.

The culture medium used in the fermentation process in accordance with the present invention usually contains nutrients for the cells or microorganisms such as digestible nitrogen sources and inorganic substances, vitamins, micro- and trace elements and other growth-promoting factors. In addition, the culture medium contains a carbon source. Various organic or inorganic substances may be used as nitrogen sources in the fermentation process in accordance with the present invention, such as nitrates, ammonium salts, yeast extract, meat extract, peptone, casein, cornsteep liquor, amino acids and urea. Typical inorganic substances that can be used in the fermentation are calcium, iron, zinc, nickel, manganese, cobalt, copper, molybdenum, and alkali salts such as chlorides, sulphates and phosphates as well as boric acid. As a carbon source, glycerol or sugar-like mono-, di-, oligo- or polysaccharides, e.g., glucose, fructose, sucrose, maltose, starch, glycogen, cellulose or substrates containing such substances, e.g., molasses, glucose syrups and fructose syrups can be used. The concentration of glucose and / or methanol in the total feed stream may vary from about 10 to about 500 g/l for each component and is preferably from about 200 to about 300 g/l. While the fermentation medium is principally an aqueous medium such medium may contain organic solvents such as alcohols, e.g. methanol, ethanol or isopropanol. Further, the fermentation medium may also be a dispersion or suspension, in which case the fermentation is suitably carried out with stirring.

For continuous operation, the cells are optionally immobilized on a solid porous carrier. Any solid porous carrier with any porosity, size and geometry conventionally used in fermentation processes and exerting no toxic effects on the particular cell or microorganism which is to be immobilized can be used for the purpose of this invention. Examples of such carriers are those made from inorganic material and having a pore diameter of from about 0.5 to about 100 μm, preferably from about 10 to about 30 μm diameter. Examples of inorganic materials are ceramics and natural minerals such as

steatite, zeolite, bentonite, silicates (glasses), aluminum silicates, aluminum oxide, magnesium aluminum silicates and magnesium aluminum oxides. Such carriers are commercially available, e.g., from Ceramtec, Marktredwitz, Germany, Schott Engineering GmbH, Mainz, Germany and others. Preferably, the carriers are spherical with a mean diameter of from about 0,2 to about 20 mm diameter. The carriers can be loaded with the living cells in a manner known per se by contacting the carrier particles with an appropriate cell culture. If desired, the carrier particles loaded with the cells can be further processed by applying a membrane-type coating layer, such as described in German Offenlegungsschrift DE 3421049. Suitably, the carrier is present in the fermentation vessel on a fixed bed. Further, the culture medium, its components and their containments, respectively are suitably sterilized prior to use if autosterilization (e.g., by methanol, ethanol, ammonia) cannot be guaranteed. Heat sterilization with steam (e.g., at 121°C and 1 bar pressure during 20 minutes) and filtration (0.2 µm) for sensitive components are preferred. Alternative sterilization methods may be applied. Media components need not necessarily be sterilized when running the process in continuous mode.

Depending on the particular cell or organism used the fermentation may be carried out at a pH between about 2 and about 11. In a preferred aspect of the invention, the fermentation process for the manufacture of phytase is carried out using the microorganism, *Hansenula polymorpha* transformed by a phytase encoding DNA sequence as described in EP 897 010, EP 897 985, or Example 11 of the present case. According to that particular aspect of the invention, the preferred carbon source is a mixture of glucose and methanol. Further, in accordance with that particular aspect of the invention, the fermentation may be carried out at a pH betweeen about 4 and 5, preferably at about pH 4.6. A preferred temperature range for carrying out such fermentation process is between about 10 and 50 °C, more preferably the fermentation temperature is about 30 °C. The aeration rate is preferably adjusted to between about 0.01 and about 1.5 volume of gas per volume of liquid with a dissolved oxygen concentration (DO) of in between 0.01 and about 500 %. A DO of 100 % denotes oxygen saturation of the solution at atmospheric pressure (1 bar) and reactor temperature. The fermentation can be carried out at a pressure of from about 0.1 to about 100 bar, preferably, the fermentation is carried out at atmospheric pressure, i.e., at about 1 bar. The dilution rate can vary from about 0.001 to about 0.5 per hour.

The invention is illustrated further by the Examples given below.

Example 1

Storage solutions for feed medium were prepared as follows:

1.1 $CaCl_2/H_3BO_3$ Solution

$CaCl_2 \cdot 2\ H_2O$	18.75	g/l
H_3BO_3	0.0125	g/l

This solution was sterilized at 121 °C for 20 minutes.

1.2 Microelements Solution

$(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$	2.5	g/l
$CuSO_4 \cdot 5\ H_2O$	0.2	g/l
$ZnSO_4 \cdot 7\ H_2O$	0.75	g/l
$MnSO_4 \cdot 5\ H_2O$	1.0	g/l
Na-EDTA	2.5	g/l

This solution was sterilized at 121 °C for 20 minutes.

1.3 Trace Elements Solution

$NiSO_4 \cdot 6\ H_2O$	0.025	g/l
$CoCl_2 \cdot 6\ H_2O$	0.025	g/l
$Na_2MoO_4 \cdot 2\ H_2O$	0.025	g/l

KJ

This solution was sterilized at 121 °C for 20 minutes.

1.4 Salts + Vitamin Solution

KH_2PO_4	50.0	g/l
$NH_4H_2PO_4$	100.0	g/l
$MgSO_4 \cdot 7\ H_2O$	45.0	g/l
$(NH_4)_2SO_4$	50.0	g/l
KCl	23.0	g/l
NaCl	5.0	g/l
vitamin solution	5.0	ml/l

(D-biotin, 600 mg/l
thiamin·HCl 200 g/l in 50 % isopropanol/water)

The vitamin solution was sterilized by filtration (0.2 µm) and added to the salt solution that was sterilized at 121 °C for 20 minutes.

1.5 Glucose Solution

770 g of D-glucose · H2O were dissolved in 480 g of water and sterilized (121 °C, 20 min) to yield 1 l solution containing 57 % (by weight) of D-glucose.

1.6 Methanol

Pure methanol was assumed to be sterile and filled into a sterilized flask.

1.7 Antifoam

A sterilized (121 °C, 20 min) solution of 10% antifoam (Struktol J 673, Schill & Seilacher, Hamburg, Germany) was provided for supply on demand by foam-control.

1.8 Base

A solution of ca. 12,5 % (by weight) of ammonia in sterile water was filled into a sterilized flask.

Example 2

A fixed bed bioreactor (1 litre) was set up following the principle illustrated in Figure 1 with individual storage flasks being provided for the solutions 1.1 to 1.8 of Example 1. The fixed bed of porous steatite spheres (4 mm diameter, pore diameter 10-30 µm, 280 pores per ml, CeramTec, Marktredwitz, Germany) was contained by a sieve plate at the top. The reactor was sterilized (121 °C, 20 min) and thereafter filled with an inoculum culture of *Hansenula polymorpha* transformed with a phytase encoding DNA as described, e.g. in EP 897 010, EP 897 985 or Example 11. Then the connection to the storage flasks was established. The inoculum culture was grown on a medium containing glycerol as a carbon source instead of glucose. The reactor was put to batch operation until all glycerol was consumed, which was determined by a rise of the dissolved oxygen concentration. Then the feed stream was turned on and the fermentation was run under process conditions as given below:

Temperature	30	°C
pH	4.6	Diluted oxygen concentration
p_{total}	10^5	N/m^2
p_{O2}	10^5	N/m^2

Dilution rate	0.0067	h^{-1}
aeration rate	100	ml/min
V_{fluid}	1190	ml^{-1}
$V_{fixed\ bed}$	950	ml^{-1}

Substrate composition as provided by storage flasks 1-8; (actual concentrations in feed stream given) :

D-glucose	305	g/l
Methanol	264	g/l
$CaCl_2/H_3BO_3$ Solution	12.2	g/l
Microelement Solution	20.9	g/l
Trace Element Solution	17.2	g/l
Salts + vitamin Solution	44.7	g/l

Analytics:

Bio-Rad Protein Assay Kit I (Bio-Rad, Glattbrugg, Switzerland) was used to determine the total protein concentration. A factor for the calculation of phytase concentration (c_{phyt}) from total protein concentration (c_{tp}) was determined as $c_{phyt} = 0.76 \cdot c_{tp}$.

To determine the biomass in the medium two samples of 1 ml were centrifuged, washed with 1 ml of water, centrifuged again, dried at 85 °C for two days and weighed.

Results:

Under the above process conditions the biomass was 59 g/l. Given a dilution rate of 0.0067 per hour the productivity was 0.078 g of phytase per litre per hour.

In a fermentation that was run fed-batch-wise the biomass was 125 g/l; the productivity, however, was calculated to 0.054 g phytase per litre per hour.

Example 3

A fermentation in analogy to Example 2 but omitting the steatite spheres (i.e., without immobilisation of the microorganism) was carried out. A nutrient and a salt and vitamin solution of the following composition were pumped into the reactor separately:

Nutrient Solution:

$NiSO_4 \cdot 6\ H_2O$	8.33	mg/l

$CoCl_2 \cdot 6\ H_2O$	8.33	mg/l
$Na_2MoO_4 \cdot 2\ H_2O$	8.33	mg/l
KJ	8.33	mg/l
$(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$	833.33	mg/l
$CuSO_4 \cdot 5\ H_2O$	66.67	mg/l
$ZnSO_4 \cdot 7\ H_2O$	250	mg/l
$MnSO_4 \cdot 5\ H_2O$	333.33	mg/l
Na-EDTA	833.33	mg/l
$CaCl_2 \cdot 2\ H_2O$	6250	mg/l
H_3BO_3	4.17	mg/l

Salts + Vitamins Solution:

KH_2PO_4	50.0	g/l
$NH_4H_2PO_4$	100.0	g/l
$MgSO_4 \cdot 7\ H_2O$	45.0	g/l
$(NH_4)_2SO_4$	50.0	g/l
KCl	23.0	g/l
NaCl	5.0	g/l
vitamin solution	5.0	ml/l

(D-biotin, 600 mg/l

thiamin·HCl 200 g/l in 50 % isopropanol/water)

The supply of these two solutions was adjusted to provide in the feed stream a concentration of 51 g/l of Nutrient Solution and 61 g/l of Salts + Vitamins Solution. The dilution rate was adjusted to 0.009 h^{-1}. The pH was kept at 4.6 by addition of 12.5 wt% ammonium hydroxide.

Furthermore, Glucose Solution as in Example 1 and methanol were fed into the reactor separately to maintain a glucose concentration of 275 g/l and a methanol concentration of 260 g/l in the feed stream.

The productivity of this fermentation was 0.088 g phytase per litre per hour. Biomass in outflow was 58 g/l.

Example 4

In a fermentation process in analogy to Example 3 but adjusting glucose concentration to 290 g/l, methanol concentration to 260 g/l, and keeping the dilution rate constant at 0.009 h^{-1}, the productivity was 0.092 g phytase per litre per hour. Biomass in outflow was 60.4 g/l.

Example 5

In a fermentation process in analogy to Example 3 but adjusting glucose concentration to 270 g/l, methanol concentration to 280 g/l, and keeping the dilution rate constant at 0.009 h^{-1}, the productivity was 0.094 g phytase per litre per hour. Biomass in outflow was 56.8 g/l.

Example 6:

Design of the amino acid sequence of consensus phytase-1

Alignment of the amino acid sequences

The alignment was calculated using the program PILEUP from the GCG Sequence Analysis Package Release 9.0 (Devereux *et al.*, 1984) with the standard parameters (gap creation penalty 12, gap extension penalty 4). The location of the gaps was refined using a text editor. Table 1 shows the sequences (see Figure 2), without the signal sequence, that were used for the performance of the alignment starting with the amino acid (aa) as mentioned in Table 1.

Table 1: Origin and vote weight of the phytase amino acid sequences used for the design of consensus phytase-1

- *phyA* from *Aspergillus terreus* 9A-1, aa 27, vote weight 0.5 (Mitchell *et al.*, 1997)
- *phyA* from *Aspergillus terreus* cbs116.46, aa 27, vote weight 0.5 (EP 897 985; Figur 1)
- *phyA* from *Aspergillus niger* var. *awamori*, aa 27, vote weight 0.33 (Piddington *et al.*, 1993)
- *phyA* from *Aspergillus niger* T213, aa 27, vote weight 0.33
- *phyA* from *Aspergillus niger* strain NRRL3135, aa 27, vote weight 0.33 (van Hartingsveldt *et al.*, 1993)
- *phyA* from *Aspergillus fumigatus* ATCC 13073, aa 26, vote weight 0.2 (Pasamontes *et al.*, 1997)
- *phyA* from *Aspergillus fumigatus* ATCC 32722, aa 26, vote weight 0.2 (EP 897 985; Figur 1)
- *phyA* from *Aspergillus fumigatus* ATCC 58128, aa 26, vote weight 0.2 (EP 897 985; Figur 1)
- *phyA* from *Aspergillus fumigatus* ATCC 26906, aa 26, vote weight 0.2 (EP 897 985; Figur 1)
- *phyA* from *Aspergillus fumigatus* ATCC 32239, aa 30, vote weight 0.2 (EP 897 985; Figur 1)
- *phy*A from *Emericella nidulans* , aa 25, vote weight 1.0 (Pasamontes *et al.*, 1997a)
- *phy*A from *Talaromyces thermophilus* ATCC 20186, aa 24, vote weight 1.0 (Pasamontes *et al.*, 1997a)
- *phy*A from *Myceliophthora thermophila*, aa 19, vote weight 1.0 (Mitchell *et al.*, 1997)

Calculation of the amino acid sequence of consensus phytase-1

Using the refined alignment as input, the consensus sequence was calculated by the program PRETTY from the GCG Sequence Analysis Package Release 9.0 (Devereux *et al.*, 1984). PRETTY prints sequences with their columns aligned and can display a consensus sequence for an alignment. A vote weight that pays regard to the similarity between the amino acid sequences of the aligned phytases was assigned to all sequences. The vote weight was set in such a way that the combined impact of all phytases from one sequence subgroup (same species, but from different strains), e. g. the amino acid sequences of all phytases from *A. fumigatus*, on the election was set one, that means that each sequence contributes with a value of 1 divided by the number of strain sequences (see Table 1). By this means, it was possible to prevent that very similar amino acid sequences, e. g. of the phytases from different *A. fumigatus* strains, dominate the calculated consensus sequence.

The program PRETTY was started with the following parameters: The plurality defining the number of votes below which there is no consensus was set on 2.0. The threshold, which determines the scoring matrix value below which an amino acid residue may not vote for a coalition of residues, was set on 2. PRETTY used the PrettyPep.Cmp consensus scoring matrix for peptides.

Ten positions of the alignment (position 46, 66, 82, 138, 162, 236, 276, 279, 280, 308; Figure 2), for which the program was not able to determine a consensus residue, were filled by hand according to the following rules: if a most frequent residue existed, this residue was chosen (138, 236, 280); if a prevalent group of similar equivalent residues occurred, the most frequent or, if not available, one residue of this group was selected (46, 66, 82, 162, 276, 308). If there was neither a prevalent residue nor a prevalent group, one of the occurring residues was chosen according to common assumptions on their influence on the protein stability (279). Eight other positions (132, 170, 204, 211, 275, 317, 384, 447; Figure 2) were not filled with the amino acid residue selected by the program but normally with amino acids that occur with the same frequency as the residues that were chosen by the program. In most cases, the slight underrating of the three *A. niger* sequences (sum of the vote weights: 0.99) was eliminated by this correction.

Conversion of the consensus phytase-1 amino acid sequence to a DNA sequence

The first 26 amino acid residues of *A. terreus* cbs116.46 phytase were used as signal peptide and, therefore, fused to the N-terminus of all consensus phytases. For this stretch, we used a special method to calculate the corresponding DNA sequence. Purvis et al

(1987) proposed that the incorporation of rare codons in a gene has an influence on the folding efficiency of the protein. The DNA sequence for the signal sequence was calculated using the approach of Purvis et al (1987) and optimized for expression in *S. cerevisiae*. For the remaining parts of the protein, we used the codon frequency table of highly expressed *S. cerevisiae* genes, obtained from the GCG program package, to translate the calculated amino acid sequence into a DNA sequence.

The resulting sequence of the *fcp* gene is shown in Figure 3.

Construction and cloning of the consensus phytase-1 gene

The calculated DNA sequence of consensus phytase-1 (*fcp*) was divided into oligonucleotides of 85 bp, alternately using the sequence of the sense and the anti-sense strand. Every oligonucleotide overlaps 20 bp with its previous and its following oligonucleotide of the opposite strand. The location of all primers, purchased from Microsynth, Balgach (Switzerland) and obtained in a PAGE-purified form, is indicated in Figure 3.

PCR-Reactions

In three PCR reactions, the synthesized oligonucleotides were composed to the entire gene. For the PCR, the High Fidelity Kit from Boehringer Mannheim (Boehringer Mannheim, Germany) and the thermo cycler The Protokol (TM) from AMS Biotechnology (Europe) Ltd. (Lugano, Switzerland) were used.

Oligonucleotides CP-1 to CP-10 (Mix 1, Figure 3) were mixed to a concentration of 0.2 pmol/µl of each oligonucleotide. A second oligonucleotide mixture (Mix 2) was prepared with CP-9 to CP-22 (0.2 pmol/µl of each oligonucleotide). Additionally, four short primers were used in the PCR reactions:

CP-a: *Eco* RI

5'-TATAT*GAATTC*ATGGGCGTGTTCGTC-3' (SEQ ID No. 1)

CP-b:

5'-TGAAAAGTTCATTGAAGGTTTC-3' (SEQ ID No. 2)

CP-c:

5'-TCTTCGAAAGCAGTACAAGTAC-3' (SEQ ID No. 3)

CP-e: *Eco* RI

5'-TATAT*GAATTC*TTAAGCGAAAC-3' (SEQ ID No. 4)

PCR reaction *a*:
10 µl Mix 1 (2.0 pmol of each oligonucleotide)
2 µl nucleotides (10 mM each nucleotide)
2 µl primer CP-a (10 pmol/µl)
2 µl primer CP-c (10 pmol/µl)
10,0 µl PCR buffer
0.75 µl polymerase mixture (2.6U)
73.25 µl H_2O

PCR reaction *b*:
10 µl Mix 2 (2.0 pmol of each oligonucleotide)
2 µl nucleotides (10 mM each nucleotide)
2 µl primer CP-b (10 pmol/µl)
2 µl primer CP-e (10 pmol/µl)
10,0 µl PCR buffer
0.75 µl polymerase mixture (2.6 U)
73.25 µl H_2O

Reaction conditions for PCR reactions *a* and *b*:

step 1	2 min - 45°C
step 2	30 sec - 72°C
step 3	30 sec - 94°C
step 4	30 sec - 52°C
step 5	1 min - 72°C

Steps 3 to 5 were repeated 40-times.

The PCR products (670 and 905 bp) were purified by an agarose gel electrophoresis (0.9% agarose) and a following gel extraction (QIAEX*II Gel Extraction Kit, Qiagen, Hilden, Germany). The purified DNA fragments were used for the PCR reaction *c*.

PCR reaction *c*:
6 µl PCR product of reaction a (≈50 ng)
6 µl PCR product of reaction b (≈50 ng)
2 µl primer CP-a (10 pmol/µl)
2 µl primer CP-e (10 pmol/µl)
10,0 µl PCR buffer
0.75 µl polymerase mixture (2.6 U)
73.25 µl H_2O

Reaction conditions for PCR reaction *c*:

*Trademark

step 1	2 min - 94°C
step 2	30 sec - 94°C
step 3	30 sec - 55°C
step 4	1 min - 72°C

Steps 2 to 4 were repeated 31-times.

The resulting PCR product (1.4 kb) was purified as mentioned above, digested with *Eco* RI, and ligated in an *Eco* RI-digested and dephosphorylated pBsk(-)-vector (Stratagene, La Jolla, CA, USA). 1 µl of the ligation mixture was used to transform *E. coli* XL-1 competent cells (Stratagene, La Jolla, CA, USA). All standard procedures were carried out as described by Sambrook *et al.* (1987). The DNA sequence of the constructed consensus phytase gene (*fcp*, Figure 3) was controlled by sequencing as known in the art.

Example 7

Design of an improved consensus phytase (consensus phytase-10) amino acid sequence

The alignments used for the design of consensus phytase-10 were calculated using the program PILEUP from the GCG Sequence Analysis Package Release 9.0 (Devereux *et al.*, 1984) with the standard parameters (gap creation penalty 12, gap extension penalty 4). The location of the gaps was refined using a text editor.

The following sequences were used for the alignment of the *Basiodiomycete* phytases starting with the amino acid (aa) mentioned in Table 2:

Table 2: Origin and vote weight of five *Basidiomycete* phytases used for the calculation of the corresponding amino acid consensus sequence (basidio)

- *phyA1* from *Paxillus involutus* NN005693, aa 21, vote weight 0.5 (WO 98/28409)
- *phyA2* from *Paxillus involutus* NN005693, aa 21, vote weight 0.5 (WO 98/28409)
- *phyA* from *Trametes pubescens* NN9343, aa 24, vote weight 1.0 (WO 98/28409)
- *phyA* from *Agrocybe pediades* NN009289, aa 19, vote weight 1.0 (WO 98/28409)
- *phyA* from *Peniophora lycii* NN006113, aa 21, vote weight 1.0 (WO 98/28409)

The alignment is shown in Figure 4.

In Table 3 the genes, which were used for the performance of the final alignment, are arranged. The first amino acid (aa) of the sequence which is used in the alignment is mentioned behind the organism's designation.

Table 3: Origin and vote weight of the phytase sequences used for the design of consensus phytase 10

- *phyA* from *Aspergillus terreus* 9A-1, aa 27, vote weight 0.5 (Mitchell *et al.*, 1997)
- *phyA* from *Aspergillus terreus* cbs116.46, aa 27, vote weight 0.5 (EP 897 985; Figur 1)
- *phyA* from *Aspergillus niger* var. *awamori*, aa 27, vote weight 0.5 (Piddington *et al.*, 1993)
- *phyA* from *Aspergillus niger* strain NRRL3135, aa 27, vote weight 0.5 (van Hartingsveldt *et al.*, 1993)
- *phyA* from *Aspergillus fumigatus* ATCC 13073, aa 26, vote weight 0.2 (Pasamontes *et al.*, 1997)
- *phyA* from *Aspergillus fumigatus* ATCC 32722, aa 26, vote weight 0.2 (EP 897 985; Figur 1)
- *phyA* from *Aspergillus fumigatus* ATCC 58128, aa 26, vote weight 0.2 (EP 897 985; Figur 1)
- *phyA* from *Aspergillus fumigatus* ATCC 26906, aa 26, vote weight 0.2 (EP 897 985; Figur 1)
- *phyA* from *Aspergillus fumigatus* ATCC 32239, aa 30, vote weight 0.2 (EP 897 985; Figur 1)
- *phy*A from *Emericella nidulans* , aa 25, vote weight 1.0 (Pasamontes *et al.*, 1997a)
- *phy*A from *Talaromyces thermophilus* ATCC 20186, aa 24, vote weight 1.0 (Pasamontes *et al.*, 1997a)
- *phy*A from *Myceliophthora thermophila*, aa 19, vote weight 1.0 (Mitchell *et al.*, 1997)
- *phyA* from *Thermomyces lanuginosa*, aa 36, vote weight 1.0 (Berka *et al.*, 1998)
- Consensus sequence of five *Basidiomycete* phytases, vote weight 1.0 (Basidio, Figure 4)

The corresponding alignment is shown in Figure 5.

Calculation of the amino acid sequence of consensus phytase-10

To improve the alignment, we combined the consensus sequence of five phytases from four different *Basidiomycetes*, called Basidio, still containing the undefined sequence positions (see Figure 4), nearly all phytase sequences used for calculation of the original consensus phytase and one new phytase sequence from the *Ascomycete Thermomyces lanuginosus* to a larger alignment.

We set plurality on 2.0 and threshold on 3. The used vote weights are listed in Table 3. The alignment and the corresponding consensus sequence are presented in Figure 5. The new consensus phytase -10 sequence has 32 different amino acids in comparison to the original consensus phytase. Positions for which the program PRETTY was not able to calculate a

consensus amino acid residue were filled according to rules mentioned in Example 6. None of the residues suggested by the program was replaced.

Furthermore, we included all Basidiomycete phytases as single amino acid sequences but assigning a vote weight of 0.2 in the alignment. The corresponding alignment is shown in Figure 7. The calculated consensus amino acid sequence (consensus phytase-11) has the following differences to the sequence of consensus phytase-10: D35X, X(K)69K, X(E)100E, A101R, Q134N, X(K)153N, X(H)190H, X(A)204S, X(E)220D, E222T, V227A, X(R)271R, H287A, X(D)288D, X(K)379K, X(I)389I, E390X, X(E)415E, X(A)416A, X(R)446L, E463A, where the numbering is as in Fig. 6.

Letter X means that the program was not able to calculate a consensus amino acid; the amino acid in parenthesis corresponds to the amino acid finally included into the consensus phytase-10.

We also checked single amino acid replacements suggested by the improved consensus phytase sequences 10 and 11 on their influence on the stability of the original consensus phytase -1. The approach is described in example 8.

Conversion of consensus phytase-10 amino acid sequence to a DNA sequence

The first 26 amino acid residues of *A. terreus* cbs116.46 phytase were used as signal peptide and, therefore, fused to the *N*-terminus of consensus phytase-10. The used procedure is further described in Example 6.

The resulting sequence of the *fcp*10 gene is shown in Figure 6.

Construction and cloning of the consensus phytase-10 gene (*fcp*10)

The calculated DNA sequence of *fcp*10 was divided into oligonucleotides of 85 bp, alternately using the sequence of the sense and the anti-sense strand. Every oligonucleotide overlaps 20 bp with its previous and its following oligonucleotide of the opposite strand. The location of all primers, purchased from Microsynth, Balgach (Switzerland) and obtained in a PAGE-purified form, is indicated in Figure 6.

PCR-Reactions

In three PCR reactions, the synthesized oligonucleotides were composed to the entire gene. For the PCR, the High Fidelity Kit from Boehringer Mannheim (Boehringer Mannheim, Mannheim, Germany) and the thermocycler The Protokol™ from AMS Biotechnology (Europe) Ltd. (Lugano, Switzerland) were used. The following oligonucleotides were used in a concentration of 0.2 pmol/ml.

Mix 1.10: CP-1, CP-2, CP-3.10, CP-4.10, CP-5.10, CP-6, CP-7.10, CP-8.10, CP-9.10, CP-10.10

Mix 2.10: CP-9.10, CP-10.10, CP-11.10, CP-12.10, CP-13.10, CP-14.10, CP-15.10, CP-16.10, CP-17.10, CP-18.10, CP-19.10, CP-20.10, CP-21.10, CP-22.10

The newly synthesized oligonucleotides are marked by number 10. The phytase contains the following 32 exchanges, which are underlined in Figure 6, in comparison to the original consensus phytase -1: Y54F, E58A, D69K, D70G, A94K, N134Q, I158V, S187A, Q188N, D197N, S204A, T214L, D220E, L234V, A238P, D246H, T251N, Y259N, E267D, E277Q, A283D, R291I, A320V, R329H, S364T, I366V, A379K, S396A, G404A, Q415E, A437G, A463E.

Four short PCR primers were used for the assembling of the oligonucleotides:

CP-a: *Eco* RI
5'-TATAT*GAATTC*ATGGGCGTGTTCGTC-3' (SEQ, ID No. 1)

CP-b:
5'-TGAAAAGTTCATTGAAGGTTTC-3' (SEQ, ID No. 2)

CP-c.10:
5'-TCTTCGAAAGCAGTACACAAAC-3' (SEQ, ID No. 5)

CP-e: *Eco* RI
5'-TATAT*GAATTC*TTAAGCGAAAC-3' (SEQ, ID No. 4)

PCR reaction *a*: 10 µl Mix 1.10 (2.0 pmol of each oligonucleotide)
2 µl nucleotides (10 mM each nucleotide)
2 µl primer CP-a (10 pmol/ml)
2 µl primer CP-c.10 (10 pmol/ml)
10,0 µl PCR buffer
0.75 µl polymerase mixture (2.6 U)
73.25 µl H_2O

PCR reaction *b*:

10 µl Mix 2.10 (2.0 pmol of each oligonucleotide)
2 µl nucleotides (10 mM each nucleotide)
2 µl primer CP-b (10 pmol/ml)
2 µl primer CP-e (10 pmol/ml)
10,0 µl PCR buffer
0.75 µl polymerase mixture (2.6 U)
73.25 µl H_2O

Reaction conditions for PCR reactions *a* and *b*:

step 1	2 min - 45 °C
step 2	30 sec - 72 °C
step 3	30 sec - 94 °C
step 4	30 sec - 52 °C
step 5	1 min - 72 °C

Steps 3 to 5 were repeated 40-times.

The PCR products (670 and 905 bp) were purified by an agarose gel electrophoresis (0.9% agarose) and a following gel extraction (QIAEX II Gel Extraction Kit, Qiagen, Hilden, Germany). The purified DNA fragments were used for the PCR reaction *c*.

PCR reaction *c*:

6 µl PCR product of reaction a(≈50 ng)
6 µl PCR product of reaction b(≈50 ng)
2 µl primer CP-a (10 pmol/ml)
2 µl primer CP-e (10 pmol/ml)
10,0 µl PCR buffer
0.75 µl polymerase mixture (2.6 U)
73.25 µl H_2O

Reaction conditions for PCR reaction *c*:

step 1	2 min - 94 °C
step 2	30 sec - 94 °C
step 3	30 sec - 55 °C
step 4	1 min - 72 °C

Steps 2 to 4 were repeated 31-times.

The resulting PCR product (1.4 kb) was purified as mentioned above, digested with *Eco* RI, and ligated in an *Eco* RI-digested and dephosphorylated pBsk(-)-vector (Stratagene, La Jolla, CA, USA). 1 µl of the ligation mixture was used to transform *E. coli* XL-1 competent cells (Stratagene, La Jolla, CA, USA). All standard procedures were carried out as described by Sambrook *et al.* (1987). The DNA sequence of the constructed gene (*fcp10*) was checked by sequencing as known in the art.

Example 8

Increasing the thermostability of consensus phytase-1 by introduction of single mutations suggested by the amino acid sequence of consensus phytase-10 and/or consensus phytase-11

In order to increase the thermostability of homologous genes, it is also possible to test the stability effect of each differing amino acid residue between the protein of interest and the calculated consensus sequence and to combine all stabilizing mutations into the protein of interest. We used the consensus phytase –1 as protein of interest and tested the effect on the protein stability of 34 amino acids, which differed between consensus phytase –1 on one hand and consensus phytases 10 and/or -11 on the other hand, by single mutation..

To construct muteins for expression in *A. niger*, *S. cerevisiae*, or *H. polymorpha*, the corresponding expression plasmid containing the consensus phytase gene was used as template for site-directed mutagenesis (see Examples 11 – 13). Mutations were introduced using the "quick exchange^{TM} site-directed mutagenesis kit" from Stratagene (La Jolla, CA, USA) following the manufacturer's protocol and using the corresponding primers. All mutations made and their corresponding primers are summarized in Table 4. Plasmids harboring the desired mutation were identified by DNA sequence analysis as known in the art.

Table 4: Primers used for site-directed mutagenesis of consensus phytases

(Exchanged bases are highlighted in bold. The introduction of a restriction site is marked above the sequence. When a restriction site is written in parenthesis, the mentioned site was destroyed by introduction of the mutation.)

mutation	Primer set
	Kpn I
Q50T	5'-CACTTGTGG*GGTACC*TACTCTCCATACTTCTC-3' (SEQ ID No. 6)
	5'-GAGAAGTATGGAGAGTA*GGTAC*CCCACAAGTG-3' (SEQ ID No. 7)
Y54F	5'-GGTCAATACTCTCCATTCTTCTCTTTGGAAG-3' (SEQ ID No. 8)
	5'-CTTCCAAAGAGAAGAATGGAGAGTATTGACC-3' (SEQ ID No. 9)
E58A	5'-CATACTTCTCTTTGGCAGACGAATCTGC-3' (SEQ ID No. 10)
	5'-GCAGATTCGTCTGCCAAAGAGAAGTATG-3' (SEQ ID No. 11)

Aat II

D69K 5'-CTCCA*GACGTC*CCAAAGGACTGTAGAGTTAC-3' (SEQ ID No. 12)
5'-GTAACTCTACAGTCCTTTGG*GACGTC*TGGAG-3' (SEQ ID No. 13)

Aat II

D70G 5'-CTCCA*GACGTC*CCAGACGGCTGTAGAGTTAC-3' (SEQ ID No. 14)
5'-GTAACTCTACAGCCGTCTGG*GACGTC*TGGAG-3' (SEQ ID No. 15)

K91A 5'-GATACCCAACTTCTTCTGCGTCTAAGGCTTACTCTG-3'
(SEQ ID No. 16)
5'-CAGAGTAAGCCTTAGACGCAGAAGAAGTTGGGTATC-3'
(SEQ ID No. 17)

Sca I

A94K 5'-CTTCTAAGTCTAAG*AAGTACT*CTGCTTTG-3' (SEQ ID No. 18)
5'-CAAAGCAG*AGTACT*TCTTAGACTTAGAAG-3'(SEQ ID No. 19)

A101R 5'-GCTTACTCTGCTTTGATTGAACGGATTCAAAAGAACGCTAC-3'

5'-GTAGCGTTCTTTTGAATCCGTTCAATCAAAGCAGAGTAAGC-3'

N134Q 5'-CCATTCGGTGAACAGCAAATGGTTAACTC-3' (SEQ ID No. 22)
5'-GAGTTAACCATTTGCTGTTCACCGAATGG-3' (SEQ ID No. 23)

Nru I

K153N 5'-GATACAAGGCTC*TCGCGA*GAAACATTGTTC –3' (SEQ ID No. 24).
5'-GGAACAATGTTTC*TCGCGA*GAGCCTTGTATC-3' (SEQ ID No. 25)

Bss HI

I158V 5'-GATTGTTCCATTCGT*GCGCGC*TTCTGGTTC-3' (SEQ ID No. 26)
5'-GAACCAGAA*GCGCGC*ACGAATGGAACAATC-3' (SEQ ID No. 27)

Bcl I

D197N 5'-CTCCAGTTATTAACG*TGATCA*TTCCAGAAGG-3' (SEQ ID No. 28)
5'-CCTTCTGGAA*TGATCA*CGTTAATAACTGGAG-3' (SEQ ID No. 29)

Apa I

S187A 5'-GGCTGACCCAG*GGGCCC*AACCACACCAAGC-3' (SEQ ID No. 30)
5'-GCTTGGTGTGGTT*GGGCCC*CTGGGTCAGCC-3' (SEQ ID No. 31)

Nco I

T214L 5'-CACTTTGGA*CCATGG*TCTTTGTACTGCTTTCG-3' (SEQ ID No. 32)
5'-CGAAAGCAGTACAAAGA*CCATGG*TCCAAAGTG-3' (SEQ ID No. 33)

Avr II

E222T 5'-GCTTTCGAAGACTCTA*CCCTAGG*TGACGACGTTG-3' (SEQ ID No. 34)
5'-CAACGTCGTCA*CCTAGGG*TAGAGTCTTCGAAAGC-3' (SEQ ID No. 35)

V227A 5'-GGTGACGACGCTGAAGCTAACTTCAC-3' (SEQ ID No. 36)
5'-GTGAAGTTAGCTTCAGCGTCGTCACC-3'(SEQ ID No. 37)

Sac II
L234V 5'-CTAACTTCA*CCGCGG*TGTTCGCTCCAG-3' (SEQ ID No. 38)
5'-CTGGAGCGAACACCGCGGTGAAGTTAG-3' (SEQ ID No. 39)

A238P 5'-GCTTTGTTCGCTCCACCTATTAGAGCTAGATTGG-3' (SEQ ID No. 40)
5'-CCAATCTAGCTCTAATAGGTGGAGCGAACAAAGC-3' (SEQ ID No. 41)

Hpa I
T251N 5'-GCCAGGT*GTTAAC*TTGACTGACGAAG-3' (SEQ ID No. 42)
5'-TTCGTCAGTCAA*GTTAAC*ACCTGGC-3' (SEQ ID No. 43)

Aat II
Y259N 5'-GACGAA*GACGTC*GTTAACTTGATGGAC-3' (SEQ ID No. 44)
5'-GTCCATCAAGTTAAC*GACGTC*TTCGTC-3' (SEQ ID No. 45)

Asp I
E267D 5'-GTCCATTC*GACACTGTC*GCTAGAACTT C-3' (SEQ ID No. 46)
5'-GAAGTTCTAGC*GACAGTGT*CGAATGGAC-3' (SEQ ID No. 47)

E277Q 5'-CTGACGCTACTCAGCTGTCTCCATTC-3' (SEQ ID No. 48)
5'-GAATGGAGACAGCTGAGTAGCGTCAG-3' (SEQ ID No. 49)

A283D 5'-GTCTCCATTCTGTGATTTGTTCACTCAC-3' (SEQ ID No. 50)
5'-GTGAGTGAACAAATCACAGAATGGAGAC-3' (SEQ ID No. 51)

Ksp I
H287A 5'-GCTTTGTTCA*CCGCGG*ACGAATGGAG-3' (SEQ ID No. 52)
5'-CTCCATTCGT*CCGCGG*TGAACAAAGC-3' (SEQ ID No. 53)

Bam HI
R291I 5'-CACGACGAAT*GGATCC*AATACGACTAC-3' (SEQ ID No. 54)
5'-GTAGTCGTATT*GGATCC*ATTCGTCGTG-3' (SEQ ID No. 55)

Bsi WI
Q292A 5'-GACGAATGGAGAG*CGTACG*ACTACTTG-3' (SEQ ID No. 56)
5'-CAAGTAGT*CGTACG*CTCTCCATTCGTC-3' (SEQ ID No. 57)

Hpa I
A320V 5'-GGTGTTGGTTTC*GTTAAC*GAATTGATTGC-3' (SEQ ID No. 58)
5'-GCAATCAATTC*GTTAAC*GAAACCAACACC-3' (SEQ ID No. 59)

(Bgl II)

R329H 5'-GCTAGATTGACT*CACTCT*CCAGTTCAAG-3' (SEQ ID No. 60)
5'-CTTGAACTGGAGA*GTGAGT*CAATCTAGC-3' (SEQ ID No. 61)

Eco RV

S364T 5'-CTCACGACAACACTAT*GATATC*TATTTTCTTC-3' (SEQ ID No. 62)
5'-GAAGAAAATA*GATATC*ATAGTGTTGTCGTGAG-3' (SEQ ID No. 63)

Nco I

I366V 5'-CGACAACT*CCATGG*TTTCTATTTTCTTCGC-3' (SEQ ID No. 64)
5'-GCGAAGAAAATAGAAA*CCATGG*AGTTGTCG-3' (SEQ ID No. 65)

Kpn I

A379K 5'-GTACAAC*GGTACC*AAGCCATTGTCTAC-3' (SEQ ID No. 66)
5'-GTAGACAATGGCTT*GGTACC*GTTGTAC-3' (SEQ ID No. 67)

S396A 5'-CTGACGGTTACGCTGCTTCTTGGAC-3' (SEQ ID No. 68
5'-GTCCAAGAAGCAGCGTAACCGTCAG-3' (SEQ ID No. 69)

G404A 5'-CTGTTCCATTCGCTGCTAGAGCTTAC-3' (SEQ ID No. 70)
5'-GTAAGCTCTAGCAGCGAATGGAACAG-3' (SEQ ID No. 71)

Q415E 5'-GATGCAATGTGAAGCTGAAAAGGAACC-3' (SEQ ID No. 72)
5'-GGTTCCTTTTCAGCTTCACATTGCATC-3' (SEQ ID No. 73)

Sal I

A437G 5'-CACGGTTGTGGT*GTCGAC*AAGTTGGG-3' (SEQ ID No. 74)
5'-CCCAACTT*GTCGAC*ACCACAACCGTG-3' (SEQ ID No. 75)

Mun I

A463E 5'-GATCTGGTGG*CAATTG*GGAGGAATGTTTCG-3' (SEQ ID No. 76)
5'-CGAAACATTCCTCC*CAATTG*CCACCAGATC-3' (SEQ ID No. 77)

and accordingly for other mutations.

The temperature optimum of the purified phytases, expressed in *Saccharomyces cerevisiae* (Example 14), was determined as outlined in Example 14. Table 5 shows the effect on the stability of consensus phytase -1 for each mutation introduced.

Table 5: Stability effect of the individual amino acid replacements in consensus phytase-1

(+ or - means a positive, respectively, negative effect on the protein stability up to 1 °C, ++ and -- means a positive, respectively, negative effect on the protein stability between 1 and

3 °C; the number 10 or 11 corresponds to the consensus phytase sequence that suggested the amino acid replacement.)

stabilizing		neutral		destabilizing	
mutation	effect	mutation	effect	mutation	effect
E58A (10)	+	D69A	±	Y54F (10)	-
D69K (11)	+	D70G (10)	±	V73I	-
D197N (10)	+	N134Q (10)	±	A94K (10)	-
T214L (10)	+ +	G186H	±	A101R (11)	-
E222T (11)	+ +	S187A (10)	±	K153N (11)	-
E267D (10)	+	T214V	±	I158V (10)	- -
R291I*	+	T251N (10)	±	G203A	- -
R329H (10)	+	Y259N (10)	±	G205S	-
S364T (10)	+ +	A283D (10)	±	A217V	-
A379K (11)	+	A320V (10)	±	V227A (11)	- -
G404A (10)	+ +	K445T	±	L234V (10)	-
		A463E (10)	±	A238P (10)	- -
				E277Q (10)	-
				H287A (11)	-
				Q292A (10)	-
				I366V (10)	-
				S396A (10)	- -
				Q415E (11)	-
				A437G (10)	- -
				E451R	- -

*: This amino acid replacement was found in another round of mutations.

We combined eight positive mutations (E58A, D197N, E267D, R291I, R329H, S364T, A379K, G404A) in consensus phytase –1 using the primers and the technique mentioned above in this example. Furthermore, the mutations Q50T and K91A were introduced which mainly influence the catalytical characteristics of the phytase (see EP 897 985 as well as Example 14). The DNA and amino acid sequence of the resulting phytase gene (consensus phytase-1-thermo[8]-Q50T-K91A) is shown in Figure 8. In this way, the temperature optimum and the melting point of the consensus phytase was increased by 7 °C (Figure 16, 17, 18).

Using the results of Table 5, we further improved the thermostability of consensus phytase 10 by the following back mutations K94A, V158I, and A396S that revealed a strong negative influence on the stability of consensus phytase -1. The resulting protein is consensus phytase-10-thermo [3]. Furthermore, we introduced the mutations Q50T and K91A which mainly influence the catalytical characteristics of consensus phytase (see EP 897 485 as well as Example 14 and Figures 15 and 16). The resulting DNA and amino acid sequence is shown in Figure 9. The optimized phytase showed a 4 °C higher temperature optimum and melting point than consensus phytase -10 (Figures 13 and 14). Furthermore, the phytase has also a strongly increased specific activity with phytate as substrate of 250 U/mg at pH 5.5 (Figure 15).

Example 9

Stabilization of the phytase of *A. fumigatus* ATCC 13073 by replacement of amino acid residues with the corresponding consensus phytase-1 and consensus phytase-10 residues

At six typical positions where the *A. fumigatus* 13073 phytase is the only or nearly the only phytase in the alignment of Figure 2 that does not contain the corresponding consensus phytase amino acid residue, the non-consensus amino acid residue was replaced by the consensus one. In a first round, the following amino acids were substituted in *A. fumigatus* 13073 phytase, containing the Q51T substitution and the signal sequence of *A. terreus* cbs.116.46 phytase (see Figure 10):

F55(28)Y, V100(73)I, F114(87)Y, A243(220)L, S265(242)P, N294(282)D.

The numbers in parentheses refer to the numbering of Figure 2.

In a second round, four of the seven stabilizing amino acid exchanges (E59A, R329H, S364T, G404A) found in the consensus phytase-10 sequence and, tested as single

mutations in consensus phytase-1 (Table 5), were additionally introduced into the *A. fumigatus* α-mutant. Furthermore, the amino acid replacement S154N, shown to reduce the protease susceptibility of the phytase, was introduced.

The mutations were introduced as described in example 8 (see Table 6) and expressed as described in example 11 to 13. The resulting *A. fumigatus* 13073 phytase variants were called a-mutant and α-mutant-E59A-S154N-R329H-S364T-G404A.

The temperature optimum (60 °C, Figure 21) and the melting point (67.0 °C, Figure 20) of the *A. fumigatus* 13073 phytase α-mutant were increased by 5 - 7°C in comparison to the values of the wild-type (temperature optimum: 55 °C, T_m: 60 °C). The five additional amino acid replacements further increased the temperature optimum by 3 °C (Figure 21).

Table 6: Mutagenesis primers for stabilization of *A. fumigatus* phytase ATCC 13073

Mutation	Primer
F55Y	5'-CACGTACTCGCCATACTTTTCGCTCGAG-3' (SEQ ID No. 78)
	5'-CTCGAGCGAAAAGTATGGCGAGTACGTG-3' (SEQ ID No. 79)
	(*Xho* I)
E58A	5'-CCATACTTTTCG*CTCGCG*GACGAGCTGTCCGTG-3' (SEQ ID NO. 80)
	5'-CACGGACAGCTCGTC*CGCGAG*CGAAAAGTAGG-3' (SEQ ID NO. 81)
V100I	5'-GTATAAGAAGCTTATTACGGCGATCCAGGCC-3' (SEQ ID No. 82)
	5'-GGCCTGGATCGCCGTAATAAGCTTCTTATAC-3' (SEQ ID No. 83)
F114Y	5'-CTTCAAGGGCAAGTACGCCTTTTTGAAGACG-3' (SEQ ID No. 84)
	5'-CGTCTTCAAAAAGGCGTACTTGCCCTTGAAG-3' (SEQ ID No. 85)
A243L	5'-CATCCGAGCTCGCCTCGAGAAGCATCTTC-3' (SEQ ID No. 86)
	5'-GAAGATGCTTCTCGAGGCGAGCTCGGATG-3' (SEQ ID No. 87)
S265P	5'-CTAATGGA TGTGTCCGTTTGATACGGTAG-3' (SEQ ID No. 88)
	5'-CTACCGTATCAAACGGACACATGTCCATTAG-3' (SEQ ID No. 89)

N294D 5'-GTGGAAGAAGTACGACTACCTTCAGTC-3' (SEQ ID No. 90)
5'-GACTGAAGGTAGTCGTACTTCTTCCAC-3' (SEQ ID No. 91)

(*Mlu* I)

R329H 5'-GCCCGGTTG*ACGCAT*TCGCCAGTGCAGG-3' (SEQ ID No. 92)
5'-CCTGCACTGGCGA*ATGCGT*CAACCGGGC-3' (SEQ ID No. 93)

Nco I

S364T 5'-CACACGACAACA*CCATGG*TTTCCATCTTC-3' (SEQ ID No. 94)
5'-GAAGATGGAAA<u>CCATGG</u>TGTTGTCGTGTG-3' (SEQ ID No. 95)

(*Bss* HI)

G404A 5'-GTGGTGCCTTTCG*CCGCGC*GAGCCTACTTC-3' (SEQ ID No. 96)
5'-GAAGTAGGCTC*GCGCGG*CGAAAGGCACCAC-3' (SEQ ID No. 97)

Example 10

Introduction of the active site amino acid residues of the *A. niger* NRRL 3135 phytase into the consensus phytase-1

We used the crystal structure of the *Aspergillus niger* NRRL 3135 phytase to define all active site amino acid residues (see Reference Example and EP 897 010). Using the alignment of Figure 2, we replaced the following active site residues and additionally the non-identical adjacent ones of the consensus phytase –1 by those of the *A. niger* phytase:

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

The new protein sequence consensus phytase -7 was backtranslated into a DNA sequence (Figure 11) as described in Example 6. The corresponding gene (*fcp7*) was generated as described in Example 6 using the following oligonucleotide mixes:

Mix 1.7: CP-1, CP-2, CP-3, CP-4.7, CP-5.7, CP-6, CP-7, CP-8.7, CP-9, CP-10.7

Mix 2.7: CP-9, CP-10.7, CP-11.7, CP-12.7, CP-13.7, CP-14.7, CP-15.7, CP-16, CP-17.7, CP-18.7, CP-19.7, CP-20, CP-21, CP-22.

The DNA sequences of the oligonucleotides are indicated in Figure 11. The newly synthesized oligonucleotides are additionally marked by number 7. After assembling of the

oligonucleotides using the same PCR primers as mentioned in Example 6, the gene was cloned into an expression vector as described in Examples 11 - 13.

The pH-profile of consensus phytase-7, purified after expression in Hansenula polymorpha, was very similar to that of A. niger NRRL 3135 phytase (see Figure 19).

Example 11

Expression of the consensus phytase genes in *Hansenula polymorpha*

The phytase expression vectors, used to transform *H. polymorpha* RB11 (Gellissen *et al.*, 1994), were constructed by inserting the *Eco* RI fragment of pBsk⁻*fcp* or variants thereof into the multiple cloning site of the *H. polymorpha* expression vector pFPMT121, which is based on an *ura3* selection marker from *S. cerevisiae*, a formate dehydrogenase (*FMD*) promoter element and a methanol oxidase (*MO*) termimator element from *H. polymorpha*. The 5' end of the *fcp* gene is fused to the *FMD* promoter, the 3' end to the *MOX* terminator (Gellissen *et al.*, 1996; EP 0299 108 B). The resulting expression vectors were designated pFPMT*fcp*, pFPMT*fcp10*, pFPMT*fcp*7.

The constructed plasmids were propagated in *E. coli*. Plasmid DNA was purified using standard state of the art procedures. The expression plasmids were transformed into the *H. polymorpha* strain RP11 deficient in orotidine-5'-phosphate decarboxylase (*ura3*) using the procedure for preparation of competent cells and for transformation of yeast as described in Gelissen *et al.* (1996). Each transformation mixture was plated on YNB (0.14% w/v Difco YNB and 0.5% ammonium sulfate) containing 2% glucose and 1.8% agar and incubated at 37 °C. After 4 to 5 days individual transformant colonies were picked and grown in the liquid medium described above for 2 days at 37 °C. Subsequently, an aliquot of this culture was used to inoculate fresh vials with YNB-medium containing 2% glucose. After seven further passages in selective medium, the expression vector is integrated into the yeast genome in multimeric form. Subsequently, mitotically stable transformants were obtained by two additional cultivation steps in 3 ml non-selective liquid medium (YPD, 2% glucose, 10 g yeast extract, and 20 g peptone). In order to obtain genetically homogeneous recombinant strains an aliquot from the last stabilization culture was plated on a selective plate. Single colonies were isolated for analysis of phytase expression in YNB containing 2% glycerol instead of glucose to derepress the *fmd* promoter. Purification of the consensus phytases was done as described in Example 12.

Example 12

Expression of the consensus phytase genes in *Saccharomyces cerevisiae* and purification of the phytases from culture supernatant

The consensus phytase genes were isolated from the corresponding Bluescript* plasmid (pBsk⁻*fcp*, pBSK⁻*fcp10*, pBsk⁻*fcp7*) and ligated into the *Eco* RI sites of the expression cassette of the *Saccharomyces cerevisiae* expression vector pYES2 (Invitrogen, San Diego, CA, USA) or subcloned between the shortened GAPFL (glyceraldhyde-3-phosphate dehydrogenase) promoter and the *pho5* terminator as described by Janes *et al.* (1990). The correct orientation of the gene was checked by PCR. Transformation of *S. cerevisiae* strains. e. g. INVSc1 (Invitrogen, San Diego, CA, USA) was done according to Hinnen *et al.* (1978). Single colonies harboring the phytase gene under the control of the GAPFL promoter were picked and cultivated in 5 ml selection medium (SD-uracil, Sherman *et al.*, 1986) at 30°C under vigorous shaking (250 rpm) for one day. The preculture was then added to 500 ml YPD medium (Sherman *et al.*, 1986) and grown under the same conditions. Induction of the *gal1* promoter was done according to the manufacturer's instructions. After four days of incubation cell broth was centrifuged (7000 rpm, GS3 rotor, 15 min, 5°C) to remove the cells and the supernatant was concentrated by way of ultrafiltration in Amicon 8400 cells (PM30 membranes) and ultrafree-15 centrifugal filter devices (Biomax-30K, Millipore, Bedford, MA, USA). The concentrate (10 ml) was desalted on a 40 ml Sephadex*G25 Superfine column (Pharmacia Biotech, Freiburg, Germany), with 10 mM sodium acetate, pH 5.0, serving as elution buffer. The desalted sample was brought to 2 M $(NH_4)_2SO_4$ and directly loaded onto a 1 ml Butyl Sepharose*4 Fast Flow hydrophobic interaction chromatography column (Pharmacia Biotech, Freiburg, Germany) which was eluted with a linear gradient from 2 M to 0 M $(NH_4)_2SO_4$ in 10 mM sodium acetate, pH 5.0. Phytase was eluted in the break-through, concentrated and loaded on a 120 ml Sephacryl S-300 gel permeation chromatography column (Pharmacia Biotech, Freiburg, Germany). Consensus phytase –1 and consensus phytase -7 eluted as a homogeneous symmetrical peak and was shown by SDS-PAGE to be approx. 95% pure.

Example 13

Expression of the consensus phytase genes in *Aspergillus niger*

The Bluescript-plasmids pBsk⁻*fcp*, pBSK⁻*fcp10*, and pBsk⁻*fcp7* were used as template for the introduction of a *Bsp* HI-site upstream of the start codon of the genes and an *Eco* RV-site downstream of the stop codon. The Expand™ High Fidelity PCR Kit (Boehringer Mannheim, Mannheim, Germany) was used with the following primers:

*Trademark

Primer Asp-1:

Bsp HI

5'-TATA*TCATGA*GCGTGTTCGTCGTGCTACTGTTC-3' (SEQ ID No. 98)

Primer Asp-2 used for cloning of *fcp* and *fcp7*:

Eco RV

3'-ACCCGACTTACAAAGCGAATT*CTATA*GATATAT-5' (SEQ ID No. 99)

Primer Asp-3 used for cloning of *fcp10*:

Eco RV

3'-ACCCTTCTTACAAAGCGAATT*CTATA*GATATAT-5' (SEQ ID No. 100)

The reaction was performed as described by the supplier. The PCR-amplified *fcp*-genes had a new *Bsp* HI site at the start codon, introduced by primer Asp-1, which resulted in a replacement of the second amino acid residue glycine by serine. Subsequently, the DNA-fragment was digested with *Bsp* HI and *Eco* RV and ligated into the *Nco* I site downstream of the glucoamylase promoter of *Aspergillus niger* (*glaA*) and the *Eco* RV site upstream of the *Aspergillus nidulans* tryptophan C terminator (*trpC*) (Mullaney *et al.*, 1985). After this cloning step, the genes were sequenced to detect possible failures introduced by PCR. The resulting expression plasmids which basically correspond to the pGLAC vector as described in Example 9 of EP 684 313 contained the orotidine-5'-phosphate decarboxylase gene (*pyr4*) of *Neurospora crassa* as a selection marker. Transformation of *Aspergillus niger* and expression of the consensus phytase genes was done as described in EP 684 313. The consensus phytases were purified as described in Example 12.

Example 14

Determination of phytase activity and of temperature optimum

Phytase activity was determined basically as described by Mitchell et al (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 900 µl 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 (1995): consensus phytase -1.101; consensus phytase -7, 1.068; consensus phytase -1 10, 1.039.

In case of pH-optimum curves, purified enzymes were diluted in 10 mM sodium acetate, pH 5.0. Incubations were started by mixing aliquots of the diluted protein with an equal volume of 1% phytic acid (≈10 mM) in a series of different buffers: 0.4 M glycine/HCl, pH 2.5; 0.4 M acetate/NaOH, pH 3.0, 3.5, 4.0, 4.5, 5.0, 5.5; 0.4 M imidazole/HCl, pH 6.0, 6.5; 0.4 M Tris/HCl pH 7.0, 7.5, 8.0, 8.5, 9.0. Control experiments showed that pH was only slightly affected by the mixing step. Incubations were performed for 15 min at 37 °C as described above.

For determinations of the substrate specificities of the phytases, phytic acid in the assay mixture was replaced by 5 mM concentrations of the respective phosphate compounds. The activity tests were performed as described above.

For determination of the temperature optimum, enzyme (100 µl) and substrate solution (100 µl) 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.

The pH-optimum of the original consensus phytase was around pH 6.0-6.5 (80 U/mg). By introduction of the Q50T mutation, the pH-optimum shifted to pH 6.0 (130 U/mg). After introduction of K91A, the pH optimum shifted one pH-unit into the acidic pH-range showing a higher specific activity between pH 2.5 and pH 6.0. That was shown for the stabilized mutants and for consensus phytase-10, too (Figure 15 and 16).

Consensus phytase-7, which was constructed to transfer the catalytic characteristics of the A. niger NRRL 3135 phytase into consensus phytase-1, had a pH-profile very similar to that of A. niger NRRL 3135 phytase (see Figure 19). The substrate specificity of consensus phytase-7 also resembled more to that of A. niger NRRL 3135 phytase than to that of consensus phytase-1.

The temperature optimum of consensus phytase-1 (71 °C) was 16-26 °C higher than the temperature optimum of the wild-type phytases (45-55 °C, Table 7) which were used to calculate the consensus sequence. The improved consensus phytase-10 showed a further increase of its temperature optimum to 80 °C (Figure 12). The temperature optimum of

the consensus phytase-1-thermo[8] phytase was found in the same range (78 °C) when using the supernatant of an overproducing *S. cerevisiae* strain. The highest temperature optimum reached of 82 °C was determined for consensus phytase-10-thermo[3]-Q50T-K91A.

Table 7: Temperature optimum and T_m-value of consensus phytase and of the phytases from *A. fumigatus*, *A. niger*, *E. nidulans*, and *M. thermophila*. The determination of the temperature optimum was performed as described in Example 14. The T_m-values were determined by differential scanning calorimetry as described in Example 15.

phytase	temperature optimum [°C]	*T*m [°C]
Consensus phytase-10-thermo[3]-Q50T-K91A	82	89.3
Consensus phytase-10-thermo[3]-Q50T	82	88.6
Consensus phytase-10	80	85.4
Consensus phytase-1-thermo[8]-Q50T	78	84.7
Consensus phytase-1-thermo[8]-Q50T-K91A	78	85.7
Consensus phytase-1	71	78.1
A. niger NRRL3135	55	63.3
A. fumigatus 13073	55	62.5
A. fumigatus 13073 α-mutant	60	67.0
A. fumigatus 13073 α-mutant (optimized)	63	-
A. terreus 9A-1	49	57.5
A. terreus cbs.116.46	45	58.5
E. nidulans	45	55.7
M. thermophila	55	n. d.
T. thermophilus	45	n. d.

Example 15

Determination of the melting point by differential scanning calorimetry (DSC)

In order to determine the unfolding temperature of the phytases, differential scanning calorimetry was applied as previously published by Lehmann et al (2000). Solutions of 50-60 mg/ml homogeneous phytase were used for the tests. A constant heating rate of 10 °C/min was applied up to 90-95 °C.

The determined melting points reflect the results obtained for the temperature optima (Table 7). The most stable consensus phytase designed is consensus phytase-10-thermo[3]-Q50T-K91A showing a melting temperature under the chosen conditions of 89.3 °C. This is 26 to 33.6 °C higher than the melting points of the wild-type phytases used.

Example 16

Transfer of basidiomycete phytase active site into consensus phytase-10-thermo[3]-Q50T-K91A

As described previously (Example 8), mutations derived from the basidiomycete phytase active site were introduced into the consensus phytase -10. The following five constructs a) to e) were prepared:

a) This construct is called consensus phytase -12, and it comprises a selected number of active site residues of the basidio consensus sequence. Its amino acid sequence (consphy12) is shown in Fig. 22 (the first 26 amino acids forms the signal peptide, amended positions are underlined);

b) a cluster of mutations (Cluster II) was transferred to the consensus phytase 10 sequence, viz.: S80Q, Y86F, S90G, K91A, S92A, K93T, A94R, Y95I;

c) analogously, another cluster of mutations (Cluster III) was transferred, viz.: T129V, E133A, Q143N, M136S, V137S, N138Q, S139A;

d) analogously, a further cluster of mutations (Cluster IV) was transferred, viz.: A168D, E171T, K172N, F173W;

e) and finally, a further cluster of mutations (Cluster V) was transferred, viz.: Q297G, S298D, G300D, Y305T.

These constructs were expressed as described in Examples 11 - 13.

References:

Akanuma, S., Yamagishi, A., Tanaka, N. & Oshima, T. (1998). Serial increase in the thermal stability of 3-isopropylmalate dehydrogenase from *Bacillus subtilis* by experimental evolution. *Prot. Sci.* 7, 698-705.

Arase, A., Yomo, T., Urabe, I., Hata, Y., Katsube, Y. & Okada, H. (1993). Stabilization of xylanase by random mutagenesis. *FEBS Lett.* 316, 123-127.

Berka, R. M., Rey, M. W., Brown, K. M., Byun, T. & Klotz, A. V. (1998). Molecular characterization and expression of a phytase gene from the thermophilic fungus Thermomyces lanuginosus. *Appl. Environ. Microbiol.* 64, 4423-4427.

Blaber, M., Lindstrom, J. D., Gassner, N., Xu, J., Heinz, D. W. & Matthews, B. W. (1993). Energetic cost and structural consequences of burying a hydroxyl group within the core of a protein determined from Ala'Ser and Val'Thr substitutions in T4 lysozyme. *Biochemistry* 32, 11363-11373.

Cosgrove, D.J. (1980) Inositol phosphates - their chemistry, biochemistry and physiology: studies in organic chemistry, chapter 4. Elsevier Scientific Publishing Company, Amsterdam, Oxford, New York.

Devereux, J., Haeberli, P.& Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12, 387-395.

Gellissen, G., Hollenberg, C. P., Janowicz, Z. A. (1994) Gene expression in methylotrophic yeasts . In: Smith, A. (ed.) Gene expression in recombinant microorganisms. Dekker, New York, 395-439.

Gellissen, G., Piontek, M., Dahlems, U., Jenzelewski, V., Gavagan, J. E., DiCosimo, R., Anton, D. I. & Janowicz, Z. A. (1996) Recombinant *Hansenula polymorpha* as a biocatalyst: coexpression of the spinach glycolate oxidase (*GO*) and the *S. cerevisiae* catalase T (*CTT1*) gene. *Appl. Microbiol. Biotechnol.* 46, 46-54.

Gerber, P. and Müller, K. (1995) Moloc molecular modeling software. *J. Comput. Aided Mol. Des.* 9, 251-268

Hinnen, A., Hicks, J. B. & Fink, G, R. (1978) Transformation of yeast. *Proc. Natl. Acad. Sci. USA* 75, 1929-1933.

Imanaka, T., Shibazaki, M. & Takagi, M. (1986). A new way of enhancing the thermostability of proteases. *Nature* 324, 695-697.

Janes, M., Meyhack, B., Zimmermann, W. & Hinnen, A. (1990) The influence of GAP promoter variants on hirudine production, average plasmid copy number and cell growth in *Saccharomyces cerevisiae*. *Curr. Genet.* 18, 97-103.

Karpusas, M., Baase, W. A., Matsumura, M. & Matthews, B. W. (1989). Hydrophobic packing in T4 lysozyme probed by cavity-filling mutants. *Proc. Natl. Acad. Sci.(USA)* 86, 8237-8241.

Lehmann, L., Kostrewa, D., Wyss, M., Brugger, R., D'Arcy, A., Pasamontes, L., van Loon, A. (2000), From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase, *Protein Engineering* 13, 49-57.

Margarit, I., Campagnoli, S., Frigerio, F., Grandi, G., Fillipis, V. D. & Fontana, A. (1992). Cumulative stabilizing effects of glycine to alanine substitutions in *Bacillus subtilis* neutral protease. *Prot. Eng.* 5, 543-550.

Matthews, B. W. (1987a). Genetic and structural analysis of the protein stability problem. *Biochemistry* 26, 6885-6888.

Matthews, B. W. (1993). Structural and genetic analysis of protein stability. *Annu. Rev. Biochem.* 62, 139-160.

Matthews, B. W., Nicholson, H. & Becktel, W. (1987). Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc. Natl. Acad. Sci. (USA)* 84, 6663-6667.

Mitchell, D. B., Vogel, K., Weimann, B. J., Pasamontes, L. & van Loon, A. P. G. M. (1997) The phytase subfamily of histidine acid phosphatases: isolation of genes for two novel phytases from the fungi *Aspergillus terreus* and *Myceliophthora thermophila*, *Microbiology* 143, 245-252.

Mullaney, E. J., Hamer, J. E., Roberti, K. A., Yelton, M. M. & Timberlake, W. E. (1985) Primary structure of the *trpC* gene from *Aspergillus nidulans*. *Mol. Gen. Genet.* 199, 37-46.

Munoz, V. & Serrano, L. (1995). Helix design, prediction and stability. *Curr. Opin. Biotechnol.* 6, 382-386.

Pace, N. C., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Prot. Sci.* 4, 2411-2423.

Pantoliano, M. W., Landner, R. C., Brian, P. N., Rollence, M. L., Wood, J. F. & Poulos, T. L. (1987). Protein engineering of subtilisin BPN': enhanced stabilization through the introduction of two cysteines to form a disulfide bond. *Biochemistry* 26, 2077-2082.

Pasamontes, L., Haiker, M., Henriquez-Huecas, M., Mitchell, D. B. & van Loon, A. P. G. M. (1997a). Cloning of the phytases from *Emericella nidulans* and the thermophilic fungus *Talaromyces thermophilus*. *Biochim. Biophys. Acta* 1353, 217-223.

Pasamontes, L., Haiker, M., Wyss, M., Tessier, M. & van Loon, A. P. G. M. (1997) Cloning, purification and characterization of a heat stable phytase from the fungus *Aspergillus fumigatus, Appl. Environ. Microbiol.* 63, 1696-1700.

Piddington, C. S., Houston, C. S., Paloheimo, M., Cantrell, M., Miettinen-Oinonen, A. Nevalainen, H., & Rambosek, J. (1993) The cloning and sequencing of the genes encoding phytase (*phy*) and pH 2.5-optimum acid phosphatase (*aph*) from *Aspergillus niger* var. *awamori. Gene* 133, 55-62.

Purvis, I. J., Bettany, A. J. E., Santiago, T. C., Coggins, J. R., Duncan, K., Eason, R. & Brown, A. J. P. (1987). The efficiency of folding of some proteins is increased by controlled rates of translation *in vivo. J. Mol. Biol.* 193, 413-417.

Risse, B., Stempfer, G., Rudolph, R., Schumacher, G. & Jaenicke, R. (1992). Characterization of the stability effect of point mutations of pyruvate oxidase from *Lactobacillus plantarum*: protection of the native state by modulating coenzyme binding and subunit interaction. *Prot. Sci.* 1, 1710-1718.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sauer, R., Hehir, K., Stearman, R., Weiss, M., Jeitler-Nilsson, A., Suchanek, E. & Pabo, C. (1986). An engineered intersubunit disulfide enhances the stability and DNA binding of the N-terminal domain of l-repressor. *Biochemistry* 25, 5992-5999.

Serrano, L., Day, A. G. & Fersht, A. R. (1993). Step-wise mutation of barnase to binase. A procedure for engineering increased stability of proteins and an experimental analysis of the evolution of protein stability. *J. Mol. Biol.* 233, 305-312.

Sheman, J. P., Finck, G. R. & Hicks, J. B. (1986) Laboratory course manual for methods in yeast genetics. Cold Spring Harbor University.

Steipe, B., Schiller, B., Plueckthun, A. & Steinbach, S. (1994). Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* 240, 188-192.

van den Burg, B., Vriend, G., Veltman, O. R., Venema & G., Eijsink, V. G. H. (1998). Engineering an enzyme to resist boiling. *Proc. Natl. Acad. Sci. (USA)* 95, 2056-2060.

Van Etten, R.L. (1982) Human prostatic acid phosphatase: a histidine phosphatase. *Ann. NY Acad. Sci.* 390, 27-50.

van Hartingsveldt, W., van Zeijl, C. M. F., Harteveld, G. M., Gouka, R. J., Suykerbuyk, M. E. G., Luiten, R. G. M., van Paridon, P. A., Selten, G. C. M., Veenstra, A. E., van Gorcom, R. F. M., & van den Hondel, C. A. M. J. J. (1993) Cloning, characterization and overexpression of the phytase-encoding gene (*phyA*) of *Aspergillus niger. Gene* 127, 87-94.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: F. Hoffmann-La Roche AG
(B) STREET: 124 Grenzacherstrasse
(C) CITY: Basle, CH-4070
(D) STATE:
(E) COUNTRY: Switzerland
(F) POSTAL CODE (ZIP):

(ii) TITLE OF INVENTION: Continuous Fermentation Process

(iii) NUMBER OF SEQUENCES: 113

(iv) CORRESPONDENCE ADDRESS
(A) NAME: GOWLING LAFLEUR HENDERSON LLP
(B) STREET: 160 ELGIN STREET, SUITE 2600
(C) CITY: OTTAWA
(D) PROVINCE: ONTARIO
(E) COUNTRY: CANADA
(F) POSTAL CODE: K1P 1C3

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 2,319,658
(B) FILING DATE: 2000-10-06

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: EP 99120289.6
(B) FILING DATE: 1999-10-11

PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: EP 00119676.5
(B) FILING DATE: 2000-09-08

(viii) ATTORNEY/AGENT INFORMATION
(A) NAME: GOWLING LAFLEUR HENDERSON LLP
(B) REFERENCE NUMBER: 08-888727CA

(ix) TELECOMMUNICATION INFORMATION
(A) TELEPHONE: 613-233-1781
(B) TELEFAX: 613-563-9869

(2) INFORMATION FOR SEQ ID NO:1
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:1:

tatatgaatt catgggcgtg ttcgtc 26

(2) INFORMATION FOR SEQ ID NO:2
(i) Sequence characteristics:
(a) Length: 22 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:2:

tgaaaagttc attgaaggtt tc 22

(2) INFORMATION FOR SEQ ID NO:3
(i) Sequence characteristics:
(a) Length: 22 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:3:

tcttcgaaag cagtacaagt ac 22

(2) INFORMATION FOR SEQ ID NO:4
(i) Sequence characteristics:
(a) Length: 22 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:4:

tatatgaatt cttaagcgaa ac 22

(2) INFORMATION FOR SEQ ID NO:5
(i) Sequence characteristics:
(a) Length: 22 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:5:

tcttcgaaag cagtacacaa ac 22

(2) INFORMATION FOR SEQ ID NO:6
(i) Sequence characteristics:
(a) Length: 32 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:6:

cacttgtggg gtacctactc tccatacttc tc 32

(2) INFORMATION FOR SEQ ID NO:7
(i) Sequence characteristics:
(a) Length: 32 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:7:

gagaagtatg gagagtaggt accccacaag tg 32

(2) INFORMATION FOR SEQ ID NO:8
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:8

ggtcaatact ctccattctt ctctttggaa g 31

(2) INFORMATION FOR SEQ ID NO:9
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:9:

cttccaaaga gaagaatgga gagtattgac c 31

(2) INFORMATION FOR SEQ ID NO:10
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:10:

catacttctc tttggcagac gaatctgc 28

(2) INFORMATION FOR SEQ ID NO:11
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type: Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:11:

gcagattcgt ctgccaaaga gaagtatg 28

(2) INFORMATION FOR SEQ ID NO:12
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:12:

ctccagacgt cccaaaggac tgtagagtta c 31

(2) INFORMATION FOR SEQ ID NO:13
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:13:

gtaactctac agtcctttgg gacgtctgga g 31

(2) INFORMATION FOR SEQ ID NO:14
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:14:

ctccagacgt cccagacggc tgtagagtta c 31

(2) INFORMATION FOR SEQ ID NO:15
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:15:

gtaactctac agccgtctgg gacgtctgga g 31

(2) INFORMATION FOR SEQ ID NO:16
(i) Sequence characteristics:
(a) Length: 36 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:16:

gatacccaac ttcttctgcg tctaaggctt actctg 36

(2) INFORMATION FOR SEQ ID NO:17
(i) Sequence characteristics:
(a) Length: 36 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:17:

cagagtaagc cttagacgca gaagaagttg ggtatc 36

(2) INFORMATION FOR SEQ ID NO:18
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:18:

cttctaagtc taagaagtac tctgctttg 29

(2) INFORMATION FOR SEQ ID NO:19
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:19:

caaagcagag tacttcttag acttagaag 29

(2) INFORMATION FOR SEQ ID NO:20
(i) Sequence characteristics:
(a) Length: 41 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:20:

gcttactctg ctttgattga acggattcaa aagaacgcta c 41

(2) INFORMATION FOR SEQ ID NO:21
(i) Sequence characteristics:
(a) Length: 41 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:21:

gtagcgttct tttgaatccg ttcaatcaaa gcagagtaag c 41

(2) INFORMATION FOR SEQ ID NO:22
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:22:

ccattcggtg aacagcaaat ggttaactc 29

(2) INFORMATION FOR SEQ ID NO:23
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:23:

gagttaacca tttgctgttc accgaatgg 29

(2) INFORMATION FOR SEQ ID NO:24
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:24:

gatacaaggc tctcgcgaga aacattgttc 30

(2) INFORMATION FOR SEQ ID NO:25
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:25:

ggaacaatgt ttctcgcgag agccttgtat c 31

(2) INFORMATION FOR SEQ ID NO:26
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:26:

gattgttcca ttcgtgcgcg cttctggttc 30

(2) INFORMATION FOR SEQ ID NO:27
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:27:

gaaccagaag cgcgcacgaa tggaacaatc 30

(2) INFORMATION FOR SEQ ID NO:28
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:28:

ctccagttat taacgtgatc attccagaag g 31

(2) INFORMATION FOR SEQ ID NO:29
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:29:

ccttctggaa tgatcacgtt aataactgga g 31

(2) INFORMATION FOR SEQ ID NO:30
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:30:

ggctgaccca ggggcccaac cacaccaagc 30

(2) INFORMATION FOR SEQ ID NO:31
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:31:

gcttggtgtg gttgggcccc tgggtcagcc 30

(2) INFORMATION FOR SEQ ID NO:32
(i) Sequence characteristics:
(a) Length: 32 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:32:

cactttggac catggtcttt gtactgcttt cg 32

(2) INFORMATION FOR SEQ ID NO:33
(i) Sequence characteristics:
(a) Length: 32 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:33:

cgaaagcagt acaaagacca tggtccaaag tg 32

(2) INFORMATION FOR SEQ ID NO:34
(i) Sequence characteristics:
(a) Length: 34 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:34:

gctttcgaag actctaccct aggtgacgac gttg 34

(2) INFORMATION FOR SEQ ID NO:35
(i) Sequence characteristics:
(a) Length: 34 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:35:

caacgtcgtc acctagggta gagtcttcga aagc 34

(2) INFORMATION FOR SEQ ID NO:36
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:36:

ggtgacgacg ctgaagctaa cttcac 26

(2) INFORMATION FOR SEQ ID NO:37
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:37:

gtgaagttag cttcagcgtc gtcacc 26

(2) INFORMATION FOR SEQ ID NO:38
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:38:

ctaacttcac cgcggtgttc gctccag 27

(2) INFORMATION FOR SEQ ID NO:39
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; Linear"

(xi) Sequence Description:SEQ ID NO:39:

ctggagcgaa caccgcggtg aagttag 27

(2) INFORMATION FOR SEQ ID NO:40
(i) Sequence characteristics:
(a) Length: 34 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:40:

gctttgttcg ctccacctat tagagctaga ttgg 34

(2) INFORMATION FOR SEQ ID NO:41
(i) Sequence characteristics:
(a) Length: 34 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:41

ccaatctagc tctaataggt ggagcgaaca aagc 34

(2) INFORMATION FOR SEQ ID NO:42
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:42:

gccaggtgtt aacttgactg acgaag 26

(2) INFORMATION FOR SEQ ID NO:43
(i) Sequence characteristics:
(a) Length: 25 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:43:

ttcgtcagtc aagttaacac ctggc 25

(2) INFORMATION FOR SEQ ID NO:44
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:44:

gacgaagacg tcgttaactt gatggac 27

(2) INFORMATION FOR SEQ ID NO:45
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:45:

gtccatcaag ttaacgacgt cttcgtc 27

(2) INFORMATION FOR SEQ ID NO:46
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:46:

gtccattcga cactgtcgct agaacttc 28

(2) INFORMATION FOR SEQ ID NO:47
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:47:

gaagttctag cgacagtgtc gaatggac 28

(2) INFORMATION FOR SEQ ID NO:48
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:48:

ctgacgctac tcagctgtct ccattc 26

(2) INFORMATION FOR SEQ ID NO:49
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:49:

gaatggagac agctgagtag cgtcag 26

(2) INFORMATION FOR SEQ ID NO:50
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:50:

gtctccattc tgtgatttgt tcactcac 28

(2) INFORMATION FOR SEQ ID NO:51
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:51:

gtgagtgaac aaatcacaga atggagac 28

(2) INFORMATION FOR SEQ ID NO:52
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:52:

gctttgttca ccgcggacga atggag 26

(2) INFORMATION FOR SEQ ID NO:53
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:53:

ctccattcgt ccgcggtgaa caaagc 26

(2) INFORMATION FOR SEQ ID NO:54
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:54:

cacgacgaat ggatccaata cgactac 27

(2) INFORMATION FOR SEQ ID NO:55
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:55:

gtagtcgtat tggatccatt cgtcgtg 27

(2) INFORMATION FOR SEQ ID NO:56
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:56:

gacgaatgga gagcgtacga ctacttg 27

(2) INFORMATION FOR SEQ ID NO:57
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:57:

caagtagtcg tacgctctcc attcgtc 27

(2) INFORMATION FOR SEQ ID NO:58
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:58:

ggtgttggtt tcgttaacga attgattgc 29

(2) INFORMATION FOR SEQ ID NO:59
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; Linear"

(xi) Sequence Description:SEQ ID NO:59:

gcaatcaatt cgttaacgaa accaacacc 29

(2) INFORMATION FOR SEQ ID NO:60
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:60:

gctagattga ctcactctcc agttcaag 28

(2) INFORMATION FOR SEQ ID NO:61
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:61:

cttgaactgg agagtgagtc aatctagc 28

(2) INFORMATION FOR SEQ ID NO:62
(i) Sequence characteristics:
(a) Length: 32 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:62:

ctcacgacaa cactatgata tctattttct tc 32

(2) INFORMATION FOR SEQ ID NO:63
(i) Sequence characteristics:
(a) Length: 32 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:63:

gaagaaaata gatatcatag tgttgtcgtg ag 32

(2) INFORMATION FOR SEQ ID NO:64
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:64:

cgacaactcc atggtttcta ttttcttcgc 30

(2) INFORMATION FOR SEQ ID NO:65
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:65:

gcgaagaaaa tagaaaccat ggagttgtcg 30

(2) INFORMATION FOR SEQ ID NO:66
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:66:

gtacaacggt accaagccat tgtctac 27

(2) INFORMATION FOR SEQ ID NO:67
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:67:

gtagacaatg gcttggtacc gttgtac 27

(2) INFORMATION FOR SEQ ID NO:68
(i) Sequence characteristics:
(a) Length: 25 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:68:

ctgacggtta cgctgcttct tggac 25

(2) INFORMATION FOR SEQ ID NO:69
(i) Sequence characteristics:
(a) Length: 25 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:69:

gtccaagaag cagcgtaacc gtcag 25

(2) INFORMATION FOR SEQ ID NO:70
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:70:

ctgttccatt cgctgctaga gcttac 26

(2) INFORMATION FOR SEQ ID NO:71
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:71:

gtaagctcta gcagcgaatg gaacag 26

(2) INFORMATION FOR SEQ ID NO:72
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:72:

gatgcaatgt gaagctgaaa aggaacc 27

(2) INFORMATION FOR SEQ ID NO:73
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; Linear"

(xi) Sequence Description:SEQ ID NO:73:

ggttcctttt cagcttcaca ttgcatc 27

(2) INFORMATION FOR SEQ ID NO:74
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:74:

cacggttgtg gtgtcgacaa gttggg 26

(2) INFORMATION FOR SEQ ID NO:75
(i) Sequence characteristics:
(a) Length: 26 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:75:

cccaacttgt cgacaccaca accgtg 26

(2) INFORMATION FOR SEQ ID NO:76
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:76:

gatctggtgg caattgggag gaatgtttcg 30

(2) INFORMATION FOR SEQ ID NO:77
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:77:

cgaaacattc ctcccaattg ccaccagatc 30

(2) INFORMATION FOR SEQ ID NO:78
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:78:

cacgtactcg ccatactttt cgctcgag 28

(2) INFORMATION FOR SEQ ID NO:79
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:79:

ctcgagcgaa aagtatggcg agtacgtg 28

(2) INFORMATION FOR SEQ ID NO:80
(i) Sequence characteristics:
(a) Length: 33 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:80:

ccatactttt cgctcgcgga cgagctgtcc gtg 33

(2) INFORMATION FOR SEQ ID NO:81
(i) Sequence characteristics:
(a) Length: 32 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:81:

cacggacagc tcgtccgcga gcgaaaagta gg 32

(2) INFORMATION FOR SEQ ID NO:82
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:82:

gtataagaag cttattacgg cgatccaggc c 31

(2) INFORMATION FOR SEQ ID NO:83
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:83:

ggcctggatc gccgtaataa gcttcttata c 31

(2) INFORMATION FOR SEQ ID NO:84
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:84:

cttcaagggc aagtacgcct ttttgaagac g 31

(2) INFORMATION FOR SEQ ID NO:85
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:85:

cgtcttcaaa aaggcgtact tgcccttgaa g 31

(2) INFORMATION FOR SEQ ID NO:86
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:86:

catccgagct cgcctcgaga agcatcttc 29

(2) INFORMATION FOR SEQ ID NO:87
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:87:

gaagatgctt ctcgaggcga gctcggatg 29

(2) INFORMATION FOR SEQ ID NO:88
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:88:

ctaatggatg tgtccgtttg atacggtag 29

(2) INFORMATION FOR SEQ ID NO:89
(i) Sequence characteristics:
(a) Length: 31 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:89:

ctaccgtatc aaacggacac atgtccatta g 31

(2) INFORMATION FOR SEQ ID NO:90
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:90:

gtggaagaag tacgactacc ttcagtc 27

(2) INFORMATION FOR SEQ ID NO:91
(i) Sequence characteristics:
(a) Length: 27 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:91:

gactgaaggt agtcgtactt cttccac 27

(2) INFORMATION FOR SEQ ID NO:92
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:92:

gcccggttga cgcattcgcc agtgcagg 28

(2) INFORMATION FOR SEQ ID NO:93
(i) Sequence characteristics:
(a) Length: 28 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence;primer"

(xi) Sequence Description:SEQ ID NO:93:

cctgcactgg cgaatgcgtc aaccgggc 28

(2) INFORMATION FOR SEQ ID NO:94
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:94:

cacacgacaa caccatggtt tccatcttc 29

(2) INFORMATION FOR SEQ ID NO:95
(i) Sequence characteristics:
(a) Length: 29 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:95:

gaagatggaa accatggtgt tgtcgtgtg 29

(2) INFORMATION FOR SEQ ID NO:96
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:96:

gtggtgcctt tcgccgcgcg agcctacttc 30

(2) INFORMATION FOR SEQ ID NO:97
(i) Sequence characteristics:
(a) Length: 30 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:97:

gaagtaggct cgcgcggcga aaggcaccac 30

(2) INFORMATION FOR SEQ ID NO:98
(i) Sequence characteristics:
(a) Length: 33 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:98:

tatatcatga gcgtgttcgt cgtgctactg ttc 33

(2) INFORMATION FOR SEQ ID NO:99
(i) Sequence characteristics:
(a) Length: 33 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:99:

acccgactta caaagcgaat tctatagata tat 33

(2) INFORMATION FOR SEQ ID NO:100
(i) Sequence characteristics:
(a) Length: 33 nucleotides
(b) Type: Nucleic Acid
(c) Strandedness: Single
(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; primer"

(xi) Sequence Description:SEQ ID NO:100:

acccttctta caaagcgaat tctatagata tat 33

(2) INFORMATION FOR SEQ ID NO:101
(i) Sequence characteristics:
(a) Length: 467 amino acids
(b) Type: Amino Acid
(d) Topology: Linear

(ii) Molecule Type: Protein

(xi) Sequence Description:SEQ ID NO:101:

```
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
 1               5                  10                  15
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
            20                  25                  30
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
        35                  40                  45
Gly Gln Tyr Ser Pro Tyr Phe Ser Leu Glu Asp Glu Ser Ala Ile Ser
    50                  55                  60
Pro Asp Val Pro Asp Asp Cys Arg Val Thr Phe Val Gln Val Leu Ser
65                  70                  75                  80
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys Ala Tyr Ser
                85                  90                  95
```

Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
100 105 110
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
115 120 125
Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
130 135 140
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
145 150 155 160
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
165 170 175
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro His Gln Ala
180 185 190
Ser Pro Val Ile Asp Val Ile Ile Pro Glu Gly Ser Gly Tyr Asn Asn
195 200 205
Thr Leu Asp His Gly Thr Cys Thr Ala Phe Glu Asp Ser Glu Leu Gly
210 215 220
Asp Asp Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro Ala Ile Arg
225 230 235 240
Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr Asp Glu Asp
245 250 255
Val Val Tyr Leu Met Asp Met Cys Pro Phe Glu Thr Val Ala Arg Thr
260 265 270
Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe Thr His Asp
275 280 285
Glu Trp Arg Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
290 295 300
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Ala
305 310 315 320
Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser Pro Val Gln Asp His Thr
325 330 335
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
340 345 350
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Ser Met Ile Ser Ile
355 360 365
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Ala Pro Leu Ser Thr Thr
370 375 380
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
385 390 395 400
Val Pro Phe Gly Ala Arg Ala Tyr Val Glu Met Met Gln Cys Gln Ala
405 410 415
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
420 425 430
Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
435 440 445
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Ala Glu
450 455 460
Cys Phe Ala
465

(2) INFORMATION FOR SEQ ID NO:102

(i) Sequence characteristics:

(a) Length: 1426 base pairs

(b) Type: Nucleic Acid

(c) Strandedness: Double

(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid
(a) Description/desc="Artificial Sequence; Consensus phytase-1"

(xi) Sequence Description:SEQ ID NO:102:

tatatgaatt catgggcgtg ttcgtcgtgc tactgtccat tgccaccttg ttcggttcca 60
catccggtac cgccttgggt cctcgtggta attctcactc ttgtgacact gttgacggtg 120
gttaccaatg tttcccagaa atttctcact tgtggggtca atactctcca tacttctctt 180
tggaagacga atctgctatt tctccagacg ttccagacga ctgtagagtt actttcgttc 240
aagttttgtc tagacacggt gctagatacc caacttcttc taagtctaag gcttactctg 300
ctttgattga agctattcaa aagaacgcta ctgctttcaa gggtaagtac gctttcttga 360
agacttacaa ctacactttg ggtgctgacg acttgactcc attcggtgaa aaccaaatgg 420
ttaactctgg tattaagttc tacagaagat acaaggcttt ggctagaaag attgttccat 480
tcattagagc ttctggttct gacagagtta ttgcttctgc tgaaaagttc attgaaggtt 540
tccaatctgc taagttggct gacccaggtt ctcaaccaca ccaagcttct ccagttattg 600
acgttattat tccagaagga tccggttaca acaacacttt ggaccacggt acttgtactg 660
ctttcgaaga ctctgaattg ggtgacgacg ttgaagctaa cttcactgct ttgttcgctc 720
cagctattag agctagattg gaagctgact tgccaggtgt tactttgact gacgaagacg 780
ttgtttactt gatggacatg tgtccattcg aaactgttgc tagaacttct gacgctactg 840
aattgtctcc attctgtgct ttgttcactc acgacgaatg gagacaatac gactacttgc 900
aatctttggg taagtactac ggttacggtg ctggtaaccc attgggtcca gctcaaggtg 960
ttggtttcgc taacgaattg attgctagat tgactagatc tccagttcaa gaccacactt 1020
ctactaacca cactttggac tctaacccag ctactttccc attgaacgct actttgtacg 1080
ctgacttctc tcacgacaac tctatgattt ctattttctt cgctttgggt ttgtacaacg 1140
gtactgctcc attgtctact acttctgttg aatctattga agaaactgac ggttactctg 1200
cttcttggac tgttccattc ggtgctagag cttacgttga aatgatgcaa tgtcaagctg 1260
aaaaggaacc attggttaga gttttggtta acgacagagt tgttccattg cacggttgtg 1320
ctgttgacaa gttgggtaga tgtaagagag acgacttcgt tgaaggtttg tctttcgcta 1380
gatctggtgg taactgggct gaatgtttcg cttaagaatt catata 1426

(2) INFORMATION FOR SEQ ID NO:103
(i) Sequence characteristics:
(a) Length: 467 amino acids
(b) Type: Amino Acid
(d) Topology: Linear

(ii) Molecule Type: Protein

(xi) Sequence Description:SEQ ID NO:103:

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
1 5 10 15
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
20 25 30
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
35 40 45
Gly Gln Tyr Ser Pro Phe Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
50 55 60
Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln Val Leu Ser
65 70 75 80
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys Lys Tyr Ser
85 90 95
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
100 105 110

Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
115 120 125
Thr Pro Phe Gly Glu Gln Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
130 135 140
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Val Arg Ala
145 150 155 160
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
165 170 175
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro His Gln Ala
180 185 190
Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly Tyr Asn Asn
195 200 205
Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser Glu Leu Gly
210 215 220
Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro Pro Ile Arg
225 230 235 240
Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr Asp Glu Asp
245 250 255
Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
260 265 270
Ser Asp Ala Thr Gln Leu Ser Pro Phe Cys Asp Leu Phe Thr His Asp
275 280 285
Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
290 295 300
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Val
305 310 315 320
Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
325 330 335
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
340 345 350
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ser Ile
355 360 365
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
370 375 380
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ala Ala Ser Trp Thr
385 390 395 400
Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Glu Ala
405 410 415
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
420 425 430
Leu His Gly Cys Gly Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
435 440 445
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu
450 455 460
Cys Phe Ala
465

(2) INFORMATION FOR SEQ ID NO:104

(i) Sequence characteristics:

(a) Length: 1426 base pairs

(b) Type: Nucleic Acid

(c) Strandedness: Double

(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid

(a) Description/desc="Artificial Sequence; Consensus phytase-1"

(xi) Sequence Description:SEQ ID NO:104:

tatatgaatt catgggcgtg ttcgtcgtgc tactgtccat tgccaccttg ttcggttcca 60
catccggtac cgccttgggt cctcgtggta attctcactc ttgtgacact gttgacggtg 120
gttaccaatg tttcccagaa atttctcact tgtggggtca atactctcca ttcttctctt 180
tggctgacga atctgctatt tctccagacg ttccaaaggg ttgtagagtt actttcgttc 240
aagttttgtc tagacacggt gctagatacc caacttcttc taagtctaag aagtactctg 300
ctttgattga agctattcaa aagaacgcta ctgctttcaa gggtaagtac gctttcttga 360
agacttacaa ctacactttg ggtgctgacg acttgactcc attcggtgaa caacaaatgg 420
ttaactctgg tattaagttc tacagaagat acaaggcttt ggctagaaag attgttccat 480
tcgttagagc ttctggttct gacagagtta ttgcttctgc tgaaaagttc attgaaggtt 540
tccaatctgc taagttggct gacccaggtg ctaacccaca ccaagcttct ccagttatta 600
acgttattat tccagaaggt gctggttaca acaacacttt ggaccacggt ttgtgtactg 660
ctttcgaaga atctgaattg ggtgacgacg ttgaagctaa cttcactgct gttttcgctc 720
cacctattag agctagattg gaagctcact tgccaggtgt taacttgact gacgaagacg 780
ttgttaactt gatggacatg tgtccattcg acactgttgc tagaacttct gacgctactc 840
aattgtctcc attctgtgac ttgttcactc acgacgaatg gattcaatac gactacttgc 900
aatctttggg taagtactac ggttacggtg ctggtaaccc attgggtcca gctcaaggtg 960
ttggtttcgt taacgaattg attgctagat tgactcactc tccagttcaa gaccacactt 1020
ctactaacca cactttggac tctaacccag ctactttccc attgaacgct actttgtacg 1080
ctgacttctc tcacgacaac actatggttt ctattttctt cgctttgggt ttgtacaacg 1140
gtactaagcc attgtctact acttctgttg aatctattga agaaactgac ggttacgctg 1200
cttcttggac tgttccattc gctgctagag cttacgttga aatgatgcaa tgtgaagctg 1260
aaaaggaacc attggttaga gttttggtta acgacagagt tgttccattg cacggttgtg 1320
gtgttgacaa gttgggtaga tgtaagagag acgacttcgt tgaaggtttg tctttcgcta 1380
gatctggtgg taactgggaa gaatgtttcg cttaagaatt catata 1426

(2) INFORMATION FOR SEQ ID NO:105

(i) Sequence characteristics:

(a) Length: 467 amino acids
(b) Type: Amino Acid
(d) Topology: Linear

(ii) Molecule Type: Protein

(xi) Sequence Description:SEQ ID NO:105:

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
1 5 10 15
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
20 25 30
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
35 40 45
Gly Thr Tyr Ser Pro Tyr Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
50 55 60
Pro Asp Val Pro Asp Asp Cys Arg Val Thr Phe Val Gln Val Leu Ser
65 70 75 80
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Ala Ser Lys Ala Tyr Ser
85 90 95
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
100 105 110
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu

Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
115 120 125
130 135 140
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
145 150 155 160
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
165 170 175
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro His Gln Ala
180 185 190
Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ser Gly Tyr Asn Asn
195 200 205
Thr Leu Asp His Gly Thr Cys Thr Ala Phe Glu Asp Ser Glu Leu Gly
210 215 220
Asp Asp Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro Ala Ile Arg
225 230 235 240
Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr Asp Glu Asp
245 250 255
Val Val Tyr Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
260 265 270
Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe Thr His Asp
275 280 285
Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
290 295 300
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Ala
305 310 315 320
Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
325 330 335
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
340 345 350
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Ile Ser Ile
355 360 365
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
370 375 380
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
385 390 395 400
Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Gln Ala
405 410 415
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
420 425 430
Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
435 440 445
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Ala Glu
450 455 460
Cys Phe Ala
465

(2) INFORMATION FOR SEQ ID NO:106

(i) Sequence characteristics:

(a) Length: 1404 base pairs

(b) Type: Nucleic Acid

(c) Strandedness: Double

(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid

(a) Description/desc="Artificial Sequence; Consensus phytase-1"

(xi) Sequence Description:SEQ ID NO:106:

atgggcgtgt tcgtcgtgct actgtccatt gccaccttgt tcggttccac atccggtacc 60
gccttgggtc ctcgtggtaa ttctcactct tgtgacactg ttgacggtgg ttaccaatgt 120
ttcccagaaa tttctcactt gtggggtacc tactctccat acttctcttt ggcagacgaa 180
tctgctattt ctccagacgt tccagacgac tgtagagtta ctttcgttca agttttgtct 240
agacacggtg ctagataccc aacttcttct gcgtctaagg cttactctgc tttgattgaa 300
gctattcaaa agaacgctac tgctttcaag ggtaagtacg ctttcttgaa gacttacaac 360
tacactttgg gtgctgacga cttgactcca ttcggtgaaa accaaatggt taactctggt 420
attaagttct acagaagata caaggctttg gctagaaaga ttgttccatt cattagagct 480
tctggttctg acagagttat tgcttctgct gaaaagttca ttgaaggttt ccaatctgct 540
aagttggctg acccaggttc tcaaccacac caagcttctc cagttattaa cgtgatcatt 600
ccagaaggat ccggttacaa caacactttg gaccacggta cttgtactgc tttcgaagac 660
tctgaattag gtgacgacgt tgaagctaac ttcactgctt tgttcgctcc agctattaga 720
gctagattgg aagctgactt gccaggtgtt actttgactg acgaagacgt tgtttacttg 780
atggacatgt gtccattcga cactgtcgct agaacttctg acgctactga attgtctcca 840
ttctgtgctt tgttcactca cgacgaatgg atccaatacg actacttgca aagcttgggt 900
aagtactacg gttacggtgc tggtaaccca ttgggtccag ctcaaggtgt tggtttcgct 960
aacgaattga ttgctagatt gactcactct ccagttcaag accacacttc tactaaccac 1020
actttggact ctaacccagc tactttccca ttgaacgcta ctttgtacgc tgacttctct 1080
cacgacaaca ctatgatatc tattttcttc gctttgggtt tgtacaacgg taccaagcca 1140
ttgtctacta cttctgttga atctattgaa gaaactgacg gttactctgc ttcttggact 1200
gttccattcg ctgctagagc ttacgttgaa atgatgcaat gtcaagctga aaaggaacca 1260
ttggttagag ttttggttaa cgacagagtt gttccattgc acggttgtgc tgttgacaag 1320
ttgggtagat gtaagagaga cgacttcgtt gaaggtttgt ctttcgctag atctggtggt 1380
aactgggctg aatgtttcgc ttaa 1404

(2) INFORMATION FOR SEQ ID NO:107

(i) Sequence characteristics:

(a) Length: 467 amino acids
(b) Type: Amino Acid
(d) Topology: Linear

(ii) Molecule Type: Protein

(xi) Sequence Description:SEQ ID NO:107:

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
1 5 10 15
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
20 25 30
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
35 40 45
Gly Thr Tyr Ser Pro Phe Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
50 55 60
Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln Val Leu Ser
65 70 75 80
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Ala Ser Lys Ala Tyr Ser
85 90 95
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
100 105 110
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu

Thr Pro Phe Gly Glu Gln Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
115 120 125
130 135 140
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
145 150 155 160
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
165 170 175
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro His Gln Ala
180 185 190
Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly Tyr Asn Asn
195 200 205
Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser Glu Leu Gly
210 215 220
Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro Pro Ile Arg
225 230 235 240
Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr Asp Glu Asp
245 250 255
Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
260 265 270
Ser Asp Ala Thr Gln Leu Ser Pro Phe Cys Asp Leu Phe Thr His Asp
275 280 285
Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
290 295 300
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Val
305 310 315 320
Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
325 330 335
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
340 345 350
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ser Ile
355 360 365
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
370 375 380
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
385 390 395 400
Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Glu Ala
405 410 415
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
420 425 430
Leu His Gly Cys Gly Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
435 440 445
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu
450 455 460
Cys Phe Ala
465

(2) INFORMATION FOR SEQ ID NO:108

(i) Sequence characteristics:

(a) Length: 1404 base pairs

(b) Type: Nucleic Acid

(c) Strandedness: Double

(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid

(a) Description/desc="Artificial Sequence; Consensus phytase-1"

(xi) Sequence Description:SEQ ID NO:108:

atgggcgtgt tcgtcgtgct actgtccatt gccaccttgt tcggttccac atccggtacc 60
gccttgggtc ctcgtggtaa ctctcactct tgtgacactg ttgacggtgg ttaccaatgt 120
ttcccagaaa tttctcactt gtggggtaca tactctccat tcttctcttt ggctgacgaa 180
tctgctattt ctccagacgt tccaaagggt tgtagagtta ctttcgttca agttttgtct 240
agacacggtg ctagataccc aacttcttct gcgtctaagg cgtactctgc tttgattgaa 300
gctattcaaa agaacgctac tgctttcaag ggtaagtacg ctttcttgaa gacttacaac 360
tacactttgg gtgctgacga cttgactcca ttcggtgaac aacaaatggt taactctggt 420
attaagttct acagaagata caaggctttg gctagaaaga ttgttccatt cattagagct 480
tctggttctg acagagttat tgcttctgct gaaaagttca ttgaaggttt ccaatctgct 540
aagttggctg acccaggtgc taacccacac caagcttctc cagttattaa cgttattatt 600
ccagaaggtg ctggttacaa caacactttg gaccacggtt tgtgtactgc tttcgaagaa 660
tctgaattgg gtgacgacgt tgaagctaac ttcactgctg ttttcgctcc accaattaga 720
gctagattgg aagctcactt gccaggtgtt aacttgactg acgaagacgt tgttaacttg 780
atggacatgt gtccattcga cactgttgct agaacttctg acgctactca attgtctcca 840
ttctgtgact tgttcactca cgacgaatgg attcaatacg actacttgca atctttgggt 900
aagtactacg gttacggtgc tggtaaccca ttgggtccag ctcaaggtgt tggtttcgtt 960
aacgaattga ttgctagatt gactcactct ccagttcaag accacacttc tactaaccac 1020
actttggact ctaacccagc tactttccca ttgaacgcta ctttgtacgc tgacttctct 1080
cacgacaaca ctatggtttc tattttcttc gctttgggtt tgtacaacgg tactaagcca 1140
ttgtctacta cttctgttga atctattgaa gaaactgacg gttactctgc ttcttggact 1200
gttccattcg ctgctagagc ttacgttgaa atgatgcaat gtgaagctga aaaggaacca 1260
ttggttagag ttttggttaa cgacagagtt gttccattgc acggttgtgg tgttgacaag 1320
ttgggtagat gtaagagaga cgacttcgtt gaaggtttgt ctttcgctag atctggtggt 1380
aactgggaag aatgtttcgc ttaa 1404

(2) INFORMATION FOR SEQ ID NO:109

(i) Sequence characteristics:

(a) Length: 467 amino acids
(b) Type: Amino Acid
(d) Topology: Linear

(ii) Molecule Type: Protein

(xi) Sequence Description:SEQ ID NO:109:

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
1 5 10 15
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn His Ser Lys Ser Cys
20 25 30
Asp Thr Val Asp Leu Gly Tyr Gln Cys Ser Pro Ala Thr Ser His Leu
35 40 45
Trp Gly Thr Tyr Ser Pro Tyr Phe Ser Leu Glu Asp Glu Leu Ser Val
50 55 60
Ser Ser Lys Leu Pro Lys Asp Cys Arg Ile Thr Leu Val Gln Val Leu
65 70 75 80
Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys Lys Tyr
85 90 95
Lys Lys Leu Ile Thr Ala Ile Gln Ala Asn Ala Thr Asp Phe Lys Gly
100 105 110
Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp

Leu Thr Pro Phe Gly Glu Gln Gln Leu Val Asn Ser Gly Ile Lys Phe
115 120 125
Tyr Gln Arg Tyr Lys Ala Leu Ala Arg Ser Val Val Pro Phe Ile Arg
130 135 140
Ala Ser Gly Ser Asp Arg Val Ile Ala Ser Gly Glu Lys Phe Ile Glu
145 150 155 160
Gly Phe Gln Gln Ala Lys Leu Ala Asp Pro Gly Ala Thr Asn Arg Ala
165 170 175
Ala Pro Ala Ile Ser Val Ile Ile Pro Glu Ser Glu Thr Phe Asn Asn
180 185 190
Thr Leu Asp His Gly Val Cys Thr Lys Phe Glu Ala Ser Gln Leu Gly
195 200 205
Asp Glu Val Ala Ala Asn Phe Thr Ala Leu Phe Ala Pro Asp Ile Arg
210 215 220
Ala Arg Leu Glu Lys His Leu Pro Gly Val Thr Leu Thr Asp Glu Asp
225 230 235 240
Val Val Ser Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
245 250 255
Ser Asp Ala Ser Gln Leu Ser Pro Phe Cys Gln Leu Phe Thr His Asn
260 265 270
Glu Trp Lys Lys Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
275 280 285
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Ile Gly Phe Thr
290 295 300
Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser Pro Val Gln Asp His Thr
305 310 315 320
Ser Thr Asn Ser Thr Leu Val Ser Asn Pro Ala Thr Phe Pro Leu Asn
325 330 335
Ala Thr Met Tyr Val Asp Phe Ser His Asp Asn Ser Met Val Ser Ile
340 345 350
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Glu Pro Leu Ser Arg Thr
355 360 365
Ser Val Glu Ser Ala Lys Glu Leu Asp Gly Tyr Ser Ala Ser Trp Val
370 375 380
Val Pro Phe Gly Ala Arg Ala Tyr Phe Glu Thr Met Gln Cys Lys Ser
385 390 395 400
Glu Lys Glu Pro Leu Val Arg Ala Leu Ile Asn Asp Arg Val Val Pro
405 410 415
Leu His Gly Cys Asp Val Asp Lys Leu Gly Arg Cys Lys Leu Asn Asp
420 425 430
Phe Val Lys Gly Leu Ser Trp Ala Arg Ser Gly Gly Asn Trp Gly Glu
435 440 445
Cys Phe Ser
450 455 460
465

(2) INFORMATION FOR SEQ ID NO:110

(i) Sequence characteristics:

(a) Length: 1404 base pairs

(b) Type: Nucleic Acid

(c) Strandedness: Double

(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid

(a) Description/desc="Artificial Sequence; Consensus phytase-1"

(xi) Sequence Description:SEQ ID NO:110:

atgggggttt tcgtcgttct attatctatc gcgactctgt tcggcagcac atcgggcact 60
gcgctgggcc cccgtggaaa tcactccaag tcctgcgata cggtagacct agggtaccag 120
tgctcccctg cgacttctca tctatggggc acgtactcgc catacttttc gctcgaggac 180
gagctgtccg tgtcgagtaa gcttcccaag gattgccgga tcaccttggt acaggtgcta 240
tcgcgccatg gagcgcggta cccaaccagc tccaagagca aaaagtataa gaagcttatt 300
acggcgatcc aggccaatgc caccgacttc aagggcaagt acgccttttt gaagacgtac 360
aactatactc tgggtgcgga tgacctcact ccctttgggg agcagcagct ggtgaactcg 420
ggcatcaagt tctaccagag gtacaaggct ctggcgcgca gtgtggtgcc gtttattcgc 480
gcctcaggct cggaccgggt tattgcttcg ggagagaagt tcatcgaggg gttccagcag 540
gcgaagctgg ctgatcctgg cgcgacgaac cgcgccgctc cggcgattag tgtgattatt 600
ccggagagcg agacgttcaa caatacgctg gaccacggtg tgtgcacgaa gtttgaggcg 660
agtcagctgg gagatgaggt tgcggccaat ttcactgcgc tctttgcacc cgacatccga 720
gctcgcctcg agaagcatct tcctggcgtg acgctgacag acgaggacgt tgtcagtcta 780
atggacatgt gtccgtttga tacggtagcg cgcaccagcg acgcaagtca gctgtcaccg 840
ttctgtcaac tcttcactca caatgagtgg aagaagtacg actaccttca gtccttgggc 900
aagtactacg gctacggcgc aggcaaccct ctgggaccgg ctcaggggat agggttcacc 960
aacgagctga ttgcccggtt gacgcgttcg ccagtgcagg accacaccag cactaactcg 1020
actctagtct ccaacccggc caccttcccg ttgaacgcta ccatgtacgt cgacttttca 1080
cacgacaaca gcatggtttc catcttcttt gcattgggcc tgtacaacgg cactgaaccc 1140
ttgtcccgga cctcggtgga aagcgccaag gaattggatg ggtattctgc atcctgggtg 1200
gtgcctttcg gcgcgcgagc ctacttcgag acgatgcaat gcaagtcgga aaaggagcct 1260
cttgttcgcg ctttgattaa tgaccgggtt gtgccactgc atggctgcga tgtggacaag 1320
ctggggcgat gcaagctgaa tgactttgtc aagggattga gttgggccag atctgggggc 1380
aactggggag agtgctttag ttga 1404

(2) INFORMATION FOR SEQ ID NO:111
(i) Sequence characteristics:
(a) Length: 467 amino acids
(b) Type: Amino Acid
(d) Topology: Linear

(ii) Molecule Type: Protein

(xi) Sequence Description:SEQ ID NO:111:

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
1 5 10 15
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
20 25 30
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
35 40 45
Gly Gln Tyr Ser Pro Tyr Phe Ser Leu Glu Asp Glu Ser Ala Ile Ser
50 55 60
Pro Asp Val Pro Asp Asp Cys Arg Val Thr Phe Val Gln Val Leu Ser
65 70 75 80
Arg His Gly Ala Arg Tyr Pro Thr Asp Ser Lys Gly Lys Lys Tyr Ser
85 90 95
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
100 105 110
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu

```
                115                 120                 125
Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
    130                 135                 140
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
145                 150                 155                 160
Ser Gly Ser Ser Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
                165                 170                 175
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro His Gln Ala
            180                 185                 190
Ser Pro Val Ile Asp Val Ile Ile Ser Glu Ala Ser Ser Tyr Asn Asn
        195                 200                 205
Thr Leu Asp Pro Gly Thr Cys Thr Ala Phe Glu Asp Ser Glu Leu Ala
    210                 215                 220
Asp Thr Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro Ala Ile Arg
225                 230                 235                 240
Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr Asp Thr Glu
                245                 250                 255
Val Thr Tyr Leu Met Asp Met Cys Ser Phe Glu Thr Val Ala Arg Thr
            260                 265                 270
Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe Thr His Asp
        275                 280                 285
Glu Trp Arg His Tyr Asp Tyr Leu Gln Ser Leu Lys Lys Tyr Tyr Gly
    290                 295                 300
His Gly Ala Gly Asn Pro Leu Gly Pro Thr Gln Gly Val Gly Phe Ala
305                 310                 315                 320
Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser Pro Val Gln Asp His Thr
                325                 330                 335
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
            340                 345                 350
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Gly Ile Ile Ser Ile
        355                 360                 365
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Ala Pro Leu Ser Thr Thr
    370                 375                 380
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ser Ala Trp Thr
385                 390                 395                 400
Val Pro Phe Ala Ser Arg Ala Tyr Val Glu Met Met Gln Cys Gln Ala
                405                 410                 415
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
            420                 425                 430
Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
        435                 440                 445
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Ala Glu
    450                 455                 460
Cys Phe Ala
465
```

(2) INFORMATION FOR SEQ ID NO:112

(i) Sequence characteristics:

(a) Length: 1426 base pairs

(b) Type: Nucleic Acid

(c) Strandedness: Double

(d) Topology: Linear

(ii) Molecule Type:Other nucleic acid

(a) Description/desc="Artificial Sequence; Consensus phytase-1"

(xi) Sequence Description:SEQ ID NO:112:

tatatgaatt catgggcgtg ttcgtcgtgc tactgtccat tgccaccttg ttcggttcca 60
catccggtac cgccttgggt cctcgtggta attctcactc ttgtgacact gttgacggtg 120
gttaccaatg tttcccagaa atttctcact tgtggggtca atactctcca tacttctctt 180
tggaagacga atctgctatt tctccagacg ttccagacga ctgtagagtt actttcgttc 240
aagttttgtc tagacacggt gctagatacc caactgactc taagggtaag aagtactctg 300
ctttgattga agctattcaa aagaacgcta ctgctttcaa gggtaagtac gctttcttga 360
agacttacaa ctacactttg ggtgctgacg acttgactcc attcggtgaa aaccaaatgg 420
ttaactctgg tattaagttc tacagaagat acaaggcttt ggctagaaag attgttccat 480
tcattagagc ttctggttct tctagagtta ttgcttctgc tgaaaagttc attgaaggtt 540
tccaatctgc taagttggct gacccaggtt ctcaaccaca ccaagcttct ccagttattg 600
acgttattat ttctgacgct tcttcttaca acaacacttt ggacccaggt acttgtactg 660
ctttcgaaga ctctgaattg gctgacactg ttgaagctaa cttcactgct ttgttcgctc 720
cagctattag agctagattg gaagctgact tgccaggtgt tactttgact gacactgaag 780
ttacttactt gatggacatg tgttctttcg aaactgttgc tagaacttct gacgctactg 840
aattgtctcc attctgtgct ttgttcactc acgacgaatg gagacactac gactacttgc 900
aatctttgaa gaagtactac ggtcacggtg ctggtaaccc attgggtcca actcaaggtg 960
ttggtttcgc taacgaattg attgctagat tgactagatc tccagttcaa gaccacactt 1020
ctactaacca cactttggac tctaacccag ctactttccc attgaacgct actttgtacg 1080
ctgacttctc tcacgacaac ggtattattt ctattttctt cgctttgggt ttgtacaacg 1140
gtactgctcc attgtctact acttctgttg aatctattga agaaactgac ggttactctt 1200
ctgcttggac tgttccattc gcttctagag cttacgttga aatgatgcaa tgtcaagctg 1260
aaaaggaacc attggttaga gttttggtta acgacagagt tgttccattg cacggttgtg 1320
ctgttgacaa gttgggtaga tgtaagagag acgacttcgt tgaaggtttg tctttcgcta 1380
gatctggtgg taactgggct gaatgtttcg cttaagaatt catata 1426

(2) INFORMATION FOR SEQ ID NO:113
(i) Sequence characteristics:
(a) Length: 467 amino acids
(b) Type: Amino Acid
(d) Topology: Linear

(ii) Molecule Type: Protein

(xi) Sequence Description:SEQ ID NO:113:

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
1 5 10 15

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
20 25 30

Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser Ser Asn Trp
35 40 45

Ser Pro Tyr Ser Pro Tyr Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
50 55 60

Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln Val Leu Gln
65 70 75 80

Arg His Gly Ala Arg Phe Pro Thr Ser Gly Ala Ala Thr Arg Ile Ser
85 90 95

Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
100 105 110

Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
115 120 125

Val Pro Phe Gly Ala Asn Gln Ser Ser Gln Ala Gly Ile Lys Phe Tyr
130 135 140

Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
145 150 155 160

Ser Gly Ser Asp Arg Val Ile Asp Ser Ala Thr Asn Trp Ile Glu Gly
165 170 175

Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro His Gln Ala
180 185 190

Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly Tyr Asn Asn
195 200 205

Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser Glu Leu Gly
210 215 220

Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro Pro Ile Arg
225 230 235 240

Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr Asp Glu Asp
245 250 255

Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
260 265 270

Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Asp Leu Phe Thr His Asp
275 280 285

Glu Trp Ile Gln Tyr Asp Tyr Leu Gly Asp Leu Asp Lys Tyr Tyr Gly
290 295 300

Thr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Val
305 310 315 320

Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
325 330 335

Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
340 345 350

Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ala Ile
355 360 365

Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr

370 375 380

Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Leu
385 390 395 400

Val Pro Phe Ser Ala Arg Met Tyr Val Glu Met Met Gln Cys Glu Ala
405 410 415

Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
420 425 430

Leu His Gly Cys Gly Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
435 440 445

Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu
450 455 460

Cys Phe Ala
465

WE CLAIM:

1. A fermentation assembly comprising:

a vessel for culturing living cells;

at least three storage flasks, wherein one of the storage flasks contains water, in fluid communication with the vessel for supply of liquids and a first transport means for transferring the liquids from the storage vessel;

individual appliances operably connected to the transport means for monitoring the supply of the contents of said storage flasks to said vessel;

a harvest flask in fluid communication with the vessel and second transport means for transferring the fermentation broth from the vessel to the harvest flask; and

a device operably connected to the first transport means for controlling and maintaining constant dilution rate in the vessel with varying rates of individual supply of liquid from the storage flasks to the vessel;

wherein

said vessel contains a fixed bed or an expanded bed or a moving bed of immobilized living cells;

the living cells are immobilized on a porous carrier;

said storage flasks comprise individual flasks for solutions of carbon, nitrogen, and mineral sources required for the growth of said cells and optimal formation of the desired reaction product;

said storage flasks comprise at least one individual flask containing a controlling agent;

said vessel further being equipped with inlet tubes from storage flasks for supply of liquids; pumps for transporting liquids from the storage flasks to vessel scales for monitoring the amount of liquids supplies to and discharged from the vessel; gas inlet

tubes and outlet tubes; a pump for discharging fermentation broth via outlet tubes to a harvest flask; a main controlling unit for overall process monitoring and steering; a controlling unit for monitoring and steering individual control systems for temperature, pH, gas pressure, vessel content and antifoarn agents; a circuit including a pump for gas supply and taking samples; a gas inlet and outlet flow control; and optionally, sterile filters and a thermostating unit.

2. The fermentation assembly of claim 1, wherein said vessel is further equipped with one or more sterile filters and thermostating unit.

3. A continuous process for the manufacture of desired proteins from cultures of living cells in which one or more nutrients and other reagents required for the growth of the living cells and the optimal production of the desired proteins are fed individually at a constant dilution rate into the assembly as defined in claim 1.

Figure 1

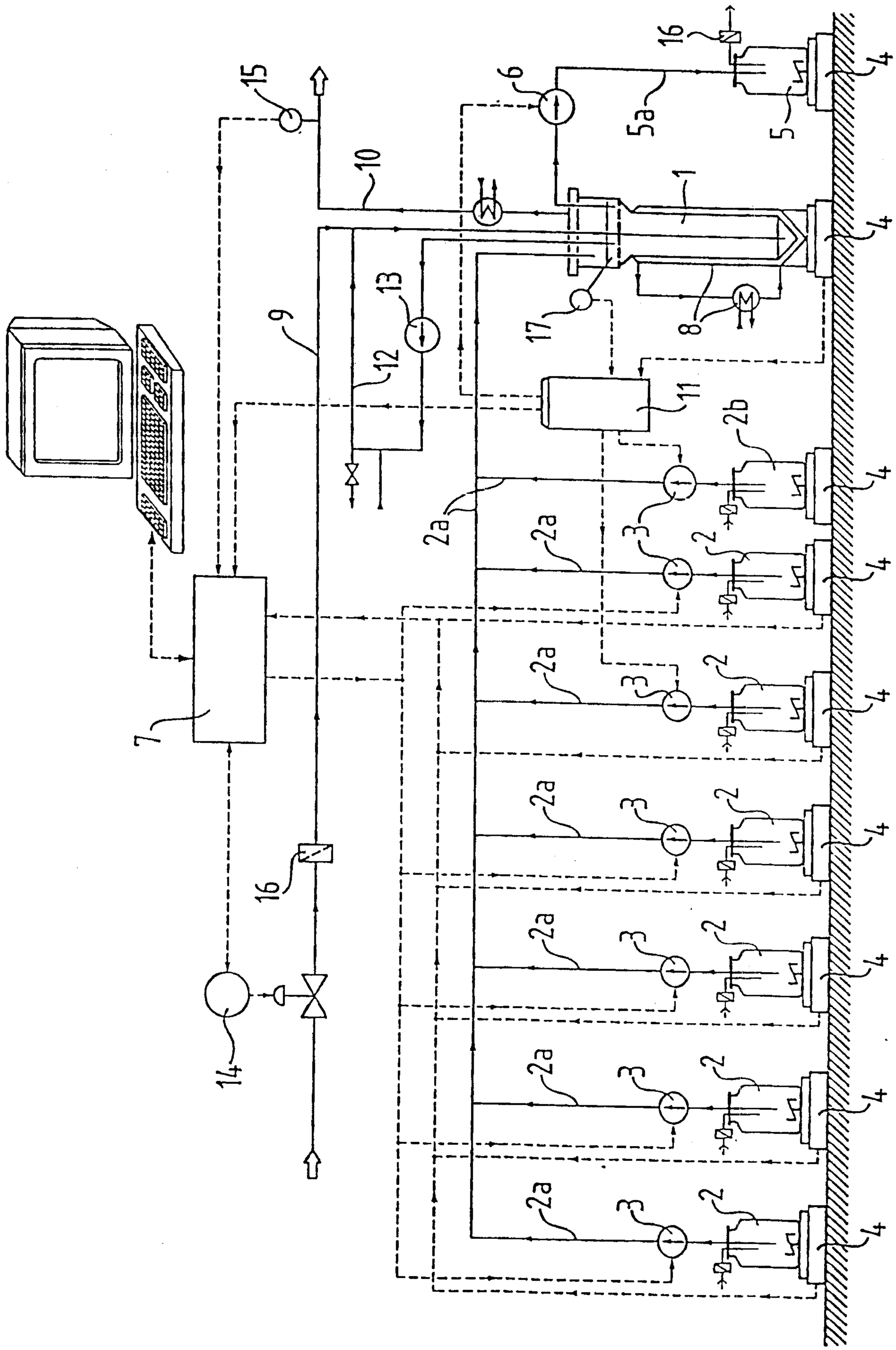


Figure 2

```
                                  1
5   50
    A. terreus 9A-1               KhsDCNSVDh GYQCFPELSH kWGlYAPYFS
    LQDESPFPlD VPEDChITFV
    A. terreus cbs                NhsDCTSVDr GYQCFPELSH kWGlYAPYFS
    LQDESPFPlD VPDDChITFV
10  A. niger var. awamori         NqsTCDTVDQ GYQCFSETSH LWGQYAPFFS
    LANESAISPD VPAGCrVTFA
    A. niger T213                 NqsSCDTVDQ GYQCFSETSH LWGQYAPFFS
    LANESVISPD VPAGCrVTFA
    A. niger NRRL3135             NqsSCDTVDQ GYQCFSETSH LWGQYAPFFS
15  LANESVISPE VPAGCrVTFA
    A. fumigatus 13073            GSkSCDTVDl GYQCsPATSH LWGQYSPFFS
    LEDElSVSSK LPKDCrITLV
    A. fumigatus 32722            GSkSCDTVDl GYQCsPATSH LWGQYSPFFS
    LEDElSVSSK LPKDCrITLV
20  A. fumigatus 58128            GSkSCDTVDl GYQCsPATSH LWGQYSPFFS
    LEDElSVSSK LPKDCrITLV
    A. fumigatus 26906            GSkSCDTVDl GYQCsPATSH LWGQYSPFFS
    LEDElSVSSK LPKDCrITLV
    A. fumigatus 32239            GSkACDTVEl GYQCsPGTSH LWGQYSPFFS
25  LEDElSVSSD LPKDCrVTFV
    E. nidulans                   QNHSCNTADG GYQCFPNVSH VWGQYSPYFS
    IEQESAISeD VPHGCeVTFV
    T. thermophilus               DSHSCNTVEG GYQCrPEISH sWGQYSPFFS
    LADQSEISPD VPQNCkITFV
30  M. thermophila                ESRPCDTpDl GFQCgTAISH FWGQYSPYFS
    VpSElDaS.. IPDDCeVTFA

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


                                  51
40  100
    A. terreus 9A-1               QVLARHGARs PThSKtKAYA AtIAAIQKSA
    TaFpGKYAFL QSYNYSLDSE
    A. terreus cbs                QVLARHGARs PTDSKtKAYA AtIAAIQKNA
    TaLpGKYAFL KSYNYSMGSE
45  A. niger var. awamori         QVLSRHGARY PTESKgKkYS ALIEEIQQNV
    TtFDGKYAFL KTYNYSLGAD
    A. niger T213                 QVLSRHGARY PTESKgKkYS ALIEEIQQNV
    TtFDGKYAFL KTYNYSLGAD
    A. niger NRRL3135             QVLSRHGARY PTDSKgKkYS ALIEEIQQNA
50  TtFDGKYAFL KTYNYSLGAD
    A. fumigatus 13073            QVLSRHGARY PTSSKsKkYK kLVTAIQaNA
    TdFKGKFAFL KTYNYTLGAD
    A. fumigatus 32722            QVLSRHGARY PTSSKsKkYK kLVTAIQaNA
    TdFKGKFAFL KTYNYTLGAD
55  A. fumigatus 58128            QVLSRHGARY PTSSKsKkYK kLVTAIQaNA
    TdFKGKFAFL KTYNYTLGAD
    A. fumigatus 26906            QVLSRHGARY PTSSKsKkYK kLVTAIQaNA
    TdFKGKFAFL KTYNYTLGAD
    A. fumigatus 32239            QVLSRHGARY PTASKsKkYK kLVTAIQKNA
60  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

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
```

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

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

Figure 3

CP-1

Eco RI M G V F V V L L S I A T L F G S T

TATAT*GAATTC*ATGGGCGTGTTCGTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCCA

5 1 ---------+---------+---------+---------+---------+---------+ 60

ATATACTTAAGTACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGT

S G T A L G P R G N S H S C D T V D G G

CATCCGGTACCGCCTTGGGTCCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTG

10 61 ---------+---------+---------+---------+---------+---------+ 120

GTAGG**CCATGGCGGAACCCAGGAGCACCATTAAGAGTGAGAACACTGTGACAACTGCCAC**

CP-2

CP-3

Y Q C F P E I S H L W G Q Y S P Y F S L

15 GTTACCAATG**TTTCCCAGAAATTTCTCACTTGTGGGGTCAATACTCTCCATACTTCTCTT**

121 ---------+---------+---------+---------+---------+---------+ 180

CAATGGTTACAAAGGGTCTTTAAAGAGTGAACACCCCAGTTATGAGAGGTATGAAGAGAA

E D E S A I S P D V P D D C R V T F V Q

20 **TGGAAGACGAATCTGCTATTTCTCCA**GACGTTCCAGACGACTGTAGAGTTACTTTCGTTC

181 ---------+---------+---------+---------+---------+---------+ 240

ACCTTCTGCTTAGAC**GATAAAGAGGTCTGCAAGGTCTGCTGACATCTCAATGAAAGCAAG**

CP-4

CP-5

25 V L S R H G A R Y P T S S K S K A Y S A

AAGTTTTGTCTAGACACGGT**GCTAGATACCCAACTTCTTCTAAGTCTAAGGCTTACTCTG**

241 ---------+---------+---------+---------+---------+---------+ 300

TTCAAAACAGATCTGTGCCACGATCTATGGGTTGAAGAAGATTCAGATTCCGAATGAGAC

L I E A I Q K N A T A F K G K Y A F L K

CTTTGATTGAAGCTATTCAAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGA

301 ---------+---------+---------+---------+---------+---------+ 360

GAAACTAACTTCGATAAGTTTTCTT**GCGATGACGAAAGTTCCCATTCATGCGAAAGAACT**

CP-6

CP-7

T Y N Y T L G A D D L T P F G E N Q M V

AGACTTACAACTACACTTTGGGTGCTGACG**ACTTGACTCCATTCGGTGAAAACCAAATGG**

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

AATTGAGACCATAATTCAAGATGTCTTCTATGTTCC**GAAACCGATCTTTCTAACAAGGTA**

CP-8

CP-9

I R A S G S D R V I A S A E K F I E G F

TCATTAGAGCTTCTGGTTCTGACAGAGTTATTGCTTCTGC**TGAAAAGTTCATTGAAGGTT**

481 ---------+---------+---------+---------+---------+---------+ 540

AGTAATCTCGAAGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAA

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

5                                                                CP-10
                                                                 CP-11
                V  I  I  P  E  G  S  G  Y  N  N  T  L  D  H  G  T  C  T  A
               ACGTTATTATTCCAGAAGGaTCcGGTTACAACAACACTTTGGACCACGGTACTTGTACTG
601 --------+---------+---------+---------+---------+---------+---------+ 660
10             TGCAATAATAAGGTCTTCCtAGgCCAATGTTGTTGTGAAACCTGGTGCCATGAACATGAC

                F  E  D  S  E  L  G  D  D  V  E  A  N  F  T  A  L  F  A  P
               CTTTCGAAGACTCTGAATTGGGTGACGACGTTGAAGCTAACTTCACTGCTTTGTTCGCTC
661 --------+---------+---------+---------+---------+---------+---------+ 720
15             GAAAGCTTCTGAGACTTAACCCACTGCTGCAACTTCGATTGAAGTGACGAAACAAGCGAG
                                                                 CP-12
                A  I  R  A  R  L  E  A  D  L  P  G  V  T  L  T  D  E  D  V
               CAGCTATTAGAGCTAGATTGGAAGCTGACTTGCCAGGTGTTACTTTGACTGACGAAGACG
20  721 --------+---------+---------+---------+---------+---------+---------+ 780
               GTCGATAATCTCGATCTAACCTTCGACTGAACGGTCCACAATGAAACTGACTGCTTCTGC
                                                                 CP-13
                V  Y  L  M  D  M  C  P  F  E  T  V  A  R  T  S  D  A  T  E
25             TTGTTTACTTGATGGACATGTGTCCATTCGAAACTGTTGCTAGAACTTCTGACGCTACTG
781 --------+---------+---------+---------+---------+---------+---------+ 840
```

AACAAATGAACTACCTGTACACAGGTAAGCTTTGACAACGATCTTGAAGACTGCGATGAC

L S P F C A L F T H D E W R Q Y D Y L Q

AATTGTCTCCATTCTGTGCTTTGTTCACTCACGACGAATGGAGACAATACGACTACTTGC

841 ---------+---------+---------+---------+---------+---------+ 900

TTAAC**AGAGGTAAGACACGAAACAAGTGAGTGCTGCTTACCTCTGTTATGCTGATGAACG**

CP-14

CP-15

S L G K Y Y G Y G A G N P L G P A Q G V

AATCTTTGGGTAAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTG

901 ---------+---------+---------+---------+---------+---------+ 960

TTAGAAACCCATTCATGATGCCAATGCCACGACCATTGGGTAACCCAGGTCGAGTTCCAC

G F A N E L I A R L T R S P V Q D H T S

TTGGTTTCGCTAACGAATTGATTGCTAGATTGACTAGATCTCCAGTTCAAGACCACACTT

961 ---------+---------+---------+---------+---------+---------+

1020

AACCAAAGCGATTGC**TTAACTAACGATCTAACTGATCTAGAGGTCAAGTTCTGGTGTGAA**

CP-16

CP-17

T N H T L D S N P A T F P L N A T L Y A

CTACTAACCACACTTTGGAC**TCTAACCCAGCTACTTTCCCATTGAACGCTACTTTGTACG**

1021 ---------+---------+---------+---------+---------+---------+

1080

GATGATTGGTGTGAAACCTGAGATTGGGTCGATGAAAGGGTAACTTGCGATGAAACATGC

```
                D  F  S  H  D  N  S  M  I  S  I  F  F  A  L  G  L  Y  N  G

               CTGACTTCTCTCACGACAACTCTATGATTTCTATTTTCTTCGCTTTGGGTTTGTACAACG

          1081 ---------+---------+---------+---------+---------+---------+
5    1140

               GACTGAAGAGAGTGCTGTTGAGATACTAAAGATAAAAGAAGCGAAACCCAAACATGTTGC

                                        CP-18

                                             CP-19

                  T  A  P  L  S  T  T  S  V  E  S  I  E  E  T  D  G  Y  S  A

10             GTACTGCTCCATTGTCTACTACTTCTGTTGAATCTATTGAAGAAACTGACGGTTACTCTG

          1141 ---------+---------+---------+---------+---------+---------+
1200

               CATGACGAGGTAACAGATGATGAAGACAACTTAGATAACTTCTTTGACTGCCAATGAGAC

15                S  W  T  V  P  F  G  A  R  A  Y  V  E  M  M  Q  C  Q  A  E

               CTTCTTGGACTGTTCCATTCGGTGCTAGAGCTTACGTTGAAATGATGCAATGTCAAGCTG

          1201 ---------+---------+---------+---------+---------+---------+
1260

               GAAGAACCTGACAAGGTAAGCCACGATCTCGAATGCAACTTTACTACGTTACAGTTCGAC

20                                                  CP-20

                                                         CP-21

                  K  E  P  L  V  R  V  L  V  N  D  R  V  V  P  L  H  G  C  A

               AAAAGGAACCATTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTG

          1261 ---------+---------+---------+---------+---------+---------+
25   1320

               TTTTCCTTGGTAACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACAC
```

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

CTGTTGACAAGTTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTA

1321 ---------+---------+---------+---------+---------+---------+ 1380

5 GACAACTGTTCAACCCATCTACATTCTCTCTGCTGAAGCAACTTC**CAAACAGAAAGCGAT**

CP-22

S G G N W A E C F A * *Eco* RI (SEQ ID No. 101)

GATCTGGTGGTAACTGGGCTGAATGTTTCGCT*TAA*GAATTCATATA (SEQ ID No. 102)

1381 ---------+---------+---------+---------+---------+------ 1426

10 **CTAGACCACCATTGACCCGACTTACAAAGCGAATTCTTAAGTATAT**

Figure 4

15 1

50

P. involutus (*phyA1*) SvP.KnTAPt FPIPeseQrn WSPYSPYFPL AeYkAPPAGC QInQVNIIQR

P. involutus (*phyA2*) SvP.RniAPK FSIPeseQrn WSPYSPYFPL AeYkAPPAGC

20 EInQVNIIQR

T. pubescens hiPlRdTSAc LdVTrDvQqs WSmYSPYFPa AtYvAPPASC QInQVHIIQR

A. pediades GgvvQaTfvQ pfFPpQiQds WAAYTPYYPV qaYtPPPkDC KItQVNIIQR

25 *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

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

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

5

401 441

P. involutus (*phyA1*) LEFCGGDrNG lCTLAkFVES QtFARsDGaG DFEKCFATSa ~

P. involutus (*phyA2*) LEFCGGDqDG lCALDkFVES QaYARsGGaG DFEKCLATTv ~

T. pubescens LAFCGADtsG vCTLDAFVES QaYARNDGEG DFEKCFAT~~ ~

10 *A. pediades* LKFCGGDmDS lCTLEAFVES QkYAREDGQG DFEKCFD~~~ ~

P. lycii LEFCGG.vDG vCeLsAFVES QtYARENGQG DFAKCgfvPs e

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

Figure 5

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

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

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.. sgggggggg EGrqeKDEeM VRVLVNDRVM TLkGCGaDEr

Basidio mvVErLxCxx xgtxxxxxxx xxxxxxxxxx VRVLVNDaVq PLEfCGgDxd

Consensus AYVEMMQC-- E--------- EG---EKEPL VRVLVNDRVV PLHGCGVDKL

Fcp10 AYVEMMQC.. EA..... .....EKEPL VRVLVNDRVV PLHGCGVDKL

451 482

A. *terreus* 9a1 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.. ..**

Figure 6

CP-1

Eco RI M G V F V V L L S I A T L F G S T 17

TATAT*GAATTC*ATGGGCGTGTTCGTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCCA

1 --------+---------+---------+---------+---------+---------+ 60

ATATACTTAAGTACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGT

S G T A L G P R G N S H S C D T V D G G 37

CATCCGGTACCGCCTTGGGTCCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTG

61 --------+---------+---------+---------+---------+---------+ 120

GTAGG**CCATGGCGGAACCCAGGAGCACCATTAAGAGTGAGAACACTGTGACAACTGCCAC**

CP-2

CP-3.10

Y Q C F P E I S H L W G Q Y S P F F S L 57

GTTACCAATG**TTTCCCAGAAATTTCTCACTTGTGGGGTCAATACTCTCCATTCTTCTCTT**

121 --------+---------+---------+---------+---------+---------+ 180

CAATGGTTACAAAGGGTCTTTAAAGAGTGAACACCCCAGTTATGAGAGGTAAGAAGAGAA

A D E S A I S P D V P K G C R V T F V Q 77

TGGCTGACGAATCTGCTATTTCTCCAGACGTTCCAAAGGGTTGTAGAGTTACTTTCGTTC

5 181 ---------+---------+---------+---------+---------+---------+ 240

ACCGACTGCTTAGAC**GATAAAGAGGTCTGCAAGGTTTCCCGACATCTCAATGAAAGCAAG**

CP-4.10

CP-5.10

V L S R H G A R Y P T S S K S K K Y S A 97

10 AAGTTTTGTCTAGACACGGT**GCTAGATACCCAACTTCTTCTAAGTCTAAGAAGTACTCTG**

241 ---------+---------+---------+---------+---------+---------+ 300

TTCAAAACAGATCTGTGCCACGATCTATGGGTTGAAGAAGATTCAGATTCTTCATGAGAC

L I E A I Q K N A T A F K G K Y A F L K 117

15 **CTTTGATTGAAGCTATTCAAAAGAACGCTACTGCTTTCAAGGGTA**AGTACGCTTTCTTGA

301 ---------+---------+---------+---------+---------+---------+ 360

GAAACTAACTTCGATAAGTTTTCTT**GCGATGACGAAAGTTCCCATTCATGCGAAAGAACT**

CP-6

CP-7.10

20 T Y N Y T L G A D D L T P F G E Q Q M V 137

AGACTTACAACTACACTTTGGGTGCTGACG**ACTTGACTCCATTCGGTGAACAACAAATGG**

361 ---------+---------+---------+---------+---------+---------+ 420

TCTGAATGTTGATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTGTTGTTTACC

25 N S G I K F Y R R Y K A L A R K I V P F 157

TTAACTCTGGTATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCAT

421 ---------+---------+---------+---------+---------+---------+ 480

AATTGAGACCATAATTCAAGATGTCTTCTATGTTC**CGAAACCGATCTTTCTAACAAGGTA**

CP-8.10

5 CP-9.10

V R A S G S D R V I A S A E K F I E G F 177

TCGTTAGAGCTTCTGGTTCTGACAGAGTTATTGCTTCTGC**TGAAAAGTTCATTGAAGGTT**

481 ---------+---------+---------+---------+---------+---------+ 540

AGCAATCTCGAAGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAA

10

Q S A K L A D P C A N P H Q A S P V I N 197

TCCAATCTGCTAAGTTGGCTGACCCAGGTGCTAACCCACACCAAGCTTCTCCAGTTATTA

541 ---------+---------+---------+---------+---------+---------+ 600

AGGTTAGACGATTCAACCGACTGGGTCCACGATTGGGTGTGGTTC**GAAGAGGTCAATAAT**

15 CP-10.10

CP-11.10

V I I P E G A G Y N N T L D H G L C T A 217

ACGTTATTATTCCAGAAGGTGCTGGTTACAACAACACTTTGGACCACGGT**TTGTGTACTG**

601 ---------+---------+---------+---------+---------+---------+ 660

20 **TGCAATAATAAGGTCTTCCACGACCAATGTTGTTGTGAAACCTGGTGCCAAACACATGAC**

F E E S E L G D D V E A N F T A V F A P 237

CTTTCGAAGAATCTGAATTGGGTGACGACGTTGAAGCTAACTTCACTGCTGTTTTCGCTC

661 ---------+---------+---------+---------+---------+---------+ 720

25 **GAAAGCTTCT**TAGACTTAACCCACTGCTGCAACTTCGATTGAAGTGACGACAAAA**GCGAG**

CP-12.10

P I R A R L E A H L P G V N L T D E D V 257

CACCTATTAGAGCTAGATTGGAAGCTCACTTGCCAGGTGTTAACTTGACTGACGAAGACG

721 ---------+---------+---------+---------+---------+---------+ 780

GTGGATAATCTCGATCTAACCTTCGAGTGAACGGTCCACAATTGAACTGACTGCTTCTGC

CP-13.10

V N L M D M C P F D T V A R T S D A T Q 277

TTGTTAACTTGATGGACATGTGTCCATTCGACACTGTTGCTAGAACTTCTGACGCTACTC

781 ---------+---------+---------+---------+---------+---------+ 840

AACAATTGAACTACCTGTACACAGGTAAGCTGTGACAACGATCTTGAAGACTGCGATGAG

L S P F C D L F T H D E W I Q Y D Y L Q 297

AATTGTCTCCATTCTGTGACTTGTTCACTCACGACGAATGGATTCAATACGACTACTTGC

841 ---------+---------+---------+---------+---------+---------+ 900

TTAAC**AGAGGTAAGACACTGAACAAGTGAGTGCTGCTTACCTAAGTTATGCTGATGAACG**

CP-14.10

CP-15.10

S L G K Y Y G Y G A G N P L G P A Q G V 317

AATCTTTGGG**TAAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTG**

901 ---------+---------+---------+---------+---------+---------+ 960

TTAGAAACCCATTCATGATGCCAATGCCACGACCATTGGGTAACCCAGGTCGAGTTCCAC

G F V N E L I A R L T H S P V Q D H T S 337

TTGGTTTCGTTAACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAAGACCACACTT

961 ---------+---------+---------+---------+---------+---------+---------+

1020

AACCAAAGCAATTGC**TTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAA**

CP-16.10

CP-17.10

T N H T L D S N P A T F P L N A T L Y A 357

CTACTAACCACACTTTGGAC**TCTAACCCAGCTACTTTCCCATTGAACGCTACTTTGTACG**

1021 ---------+---------+---------+---------+---------+---------+---------+

1080

GATGATTGGTGTGAAACCTGAGATTGGGTCGATGAAAGGGTAACTTGCGATGAAACATGC

D F S H D N T M V S I F F A L G L Y N G 377

CTGACTTCTCTCACGACAACACTATGGTTTCTATTTTCTTCGCTTTGGGTTTGTACAACG

1081 ---------+---------+---------+---------+---------+---------+---------+

1140

GACTGAAGAGAGTGCTGTTGTGATA**CCAAAGATAAAAGAAGCGAAACCCAAACATGTTGC**

CP-18.10

CP-19.10

T K P L S T T S V E S I E E T D G Y A A 397

GTACTAAGCCATTGTCTACTACTTCTGTTG**AATCTATTGAAGAAACTGACGGTTACGCTG**

1141 ---------+---------+---------+---------+---------+---------+---------+

1200

CATGATTCGGTAACAGATGATGAAGACAACTTAGATAACTTCTTTGACTGCCAATGCGAC

S W T V P F A A R A Y V E M M Q C E A E 417

CTTCTTGGACTGTTCCATTCGCTGCTAGAGCTTACGTTGAAATGATGCAATGTGAAGCTG

1201 ---------+---------+---------+---------+---------+---------+

1260

GAAGAACCTGACAAGGTAAGCCACGATCTCGAATG**CAACTTTACTACGTTACACTTCGAC**

CP-20.10

K E P L V R V L V N D R V V P L H G C G 437

AAAAGGAACCATTGGTTAGAGTTTTGGTTAACGACAGAG**TTGTTCCATTGCACGGTTGTG**

1261 ---------+---------+---------+---------+---------+---------+

1320

TTTTCCTTGGTAACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACAC

V D K L G R C K R D D F V E G L S F A R 457

GTGTTGACAAGTTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTA

1321 ---------+---------+---------+---------+---------+---------+

1380

CACAACTGTTCAACCCATCTACATTCTCTCTGCTGAAGCAACTTC**CAAACAGAAAGCGAT**

CP-22.10

S G G N W E E C F A * *Eco* RI 467 (SEQ ID No. 103)

GATCTGGTGGTAACTGGGAAGAATGTTTCGCT*TAA*GAATTCATATA (SEQ ID No. 104)

1381 ---------+---------+---------+---------+------ 1426

CTAGACCACCATTGACCCTTCTTACAAAGCGAATTCTTAAGTATAT

Figure 7

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

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

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

A. pediades mvtErL1CQr 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~
5	*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~~~	~~
10	*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~~	~~
15	*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~
20	*T. lanuginosa*	GRCRrDEWIK	GLTFARqG.. GHWDrCF~~~	~~
	M. thermophila	GmCtLErFIE	SMAFARGN.. GKWDlCFA~~	~~
	Consensus Seq. 11	**GRCKLDDFVE**	**GLSFARSG-- GNWAECFA--**	**--**

Figure 8

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

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

                 F  P  E  I  S  H  L  W  G  T  Y  S  P  Y  F  S  L  A  D  E
    60
                 TTCCCAGAAATTTCTCACTTGTGGGGTACCTACTCTCCATACTTCTCTTTGGCAGACGAA
20       121 ---+---------+---------+---------+---------+---------+---------+------
    180
                 AAGGGTCTTTAAAGAGTGAACACCCCATGGATGAGAGGTATGAAGAGAAACCGTCTGCTT

                 S  A  I  S  P  D  V  P  D  D  C  R  V  T  F  V  Q  V  L  S
25  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

5               AGACACGGTGCTAGATACCCAACTTCTTCTGCGTCTAAGGCTTACTCTGCTTTGATTGAA
          241 ---+---------+---------+---------+---------+---------+---------+------
    300

                TCTGTGCCACGATCTATGGGTTGAAGAAGACGCAGATTCCGAATGAGACGAAACTAACTT

10              A  I  Q  K  N  A  T  A  F  K  G  K  Y  A  F  L  K  T  Y  N
    120

                GCTATTCAAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGAAGACTTACAAC
          301 ---+---------+---------+---------+---------+---------+---------+------
    360

15              CGATAAGTTTTCTTGCGATGACGAAAGTTCCCATTCATGCGAAAGAACTTCTGAATGTTG

                Y  T  L  G  A  D  D  L  T  P  F  G  E  N  Q  M  V  N  S  G
    140

                TACACTTTGGGTGCTGACGACTTGACTCCATTCGGTGAAAACCAAATGGTTAACTCTGGT

20        361 ---+---------+---------+---------+---------+---------+---------+------
    420

                ATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACCAATTGAGACCA

                I  K  F  Y  R  R  Y  K  A  L  A  R  K  I  V  P  F  I  R  A
25  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

5               TCTGGTTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTTTCCAATCTGCT

        481 ---+---------+---------+---------+---------+---------+---------+------
    540

                AGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAAAGGTTAGACGA

10              K  L  A  D  P  G  S  Q  P  H  Q  A  S  P  V  I  N  V  I  I
    200

                AAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTAACGTGATCATT

        541 ---+---------+---------+---------+---------+---------+---------+------
    600

15              TTCAACCGACTGGGTCCAAGAGTTGGTGTGGTTCGAAGAGGTCAATAATTGCACTAGTAA

                P  E  G  S  G  Y  N  N  T  L  D  H  G  T  C  T  A  F  E  D
    220

                CCAGAAGGATCCGGTTACAACAACACTTTGGACCACGGTACTTGTACTGCTTTCGAAGAC

20          601 ---+---------+---------+---------+---------+---------+---------+------
    660

                GGTCTTCCTAGGCCAATGTTGTTGTGAAACCTGGTGCCATGAACATGACGAAAGCTTCTG

                S  E  L  G  D  D  V  E  A  N  F  T  A  L  F  A  P  A  I  R
25  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

5               GCTAGATTGGAAGCTGACTTGCCAGGTGTTACTTTGACTGACGAAGACGTTGTTTACTTG
        721 ---+---------+---------+---------+---------+---------+---------+------
    780

                CGATCTAACCTTCGACTGAACGGTCCACAATGAAACTGACTGCTTCTGCAACAAATGAAC

10              M  D  M  C  P  F  D  T  V  A  R  T  S  D  A  T  E  L  S  P
    280

                ATGGACATGTGTCCATTCGACACTGTCGCTAGAACTTCTGACGCTACTGAATTGTCTCCA
        781 ---+---------+---------+---------+---------+---------+---------+------
    840

15              TACCTGTACACAGGTAAGCTGTGACAGCGATCTTGAAGACTGCGATGACTTAACAGAGGT

                F  C  A  L  F  T  H  D  E  W  I  Q  Y  D  Y  L  Q  S  L  G
    300

                TTCTGTGCTTTGTTCACTCACGACGAATGGATCCAATACGACTACTTGCAAAGCTTGGGT

20          841 ---+---------+---------+---------+---------+---------+---------+------
    900

                AAGACACGAAACAAGTGAGTGCTGCTTACCTAGGTTATGCTGATGAACGTTTCGAACCCA

                K  Y  Y  G  Y  G  A  G  N  P  L  G  P  A  Q  G  V  G  F  A
25  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

5               AACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAAGACCACACTTCTACTAACCAC

      961 ---+---------+---------+---------+---------+---------+---------+-----
    1020

                TTGCTTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAAGATGATTGGTG

10              T  L  D  S  N  P  A  T  F  P  L  N  A  T  L  Y  A  D  F  S
    360

                ACTTTGGACTCTAACCCAGCTACTTTCCCATTGAACGCTACTTTGTACGCTGACTTCTCT

     1021 ---+---------+---------+---------+---------+---------+---------+-----
    1080

15              TGAAACCTGAGATTGGGTCGATGAAAGGGTAACTTGCGATGAAACATGCGACTGAAGAGA

                H  D  N  T  M  I  S  I  F  F  A  L  G  L  Y  N  G  T  K  P
    380

                CACGACAACACTATGATATCTATTTTCTTCGCTTTGGGTTTGTACAACGGTACCAAGCCA

20       1081 ---+---------+---------+---------+---------+---------+---------+-----
    1140

                GTGCTGTTGTGATACTATAGATAAAAGAAGCGAAACCCAAACATGTTGCCATGGTTCGGT

                L  S  T  T  S  V  E  S  I  E  E  T  D  G  Y  S  A  S  W  T
25  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

TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTAGATCTGGTGGT

1321 ---+---------+---------+---------+---------+---------+---------+------

1380

AACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGATCTAGACCACCA

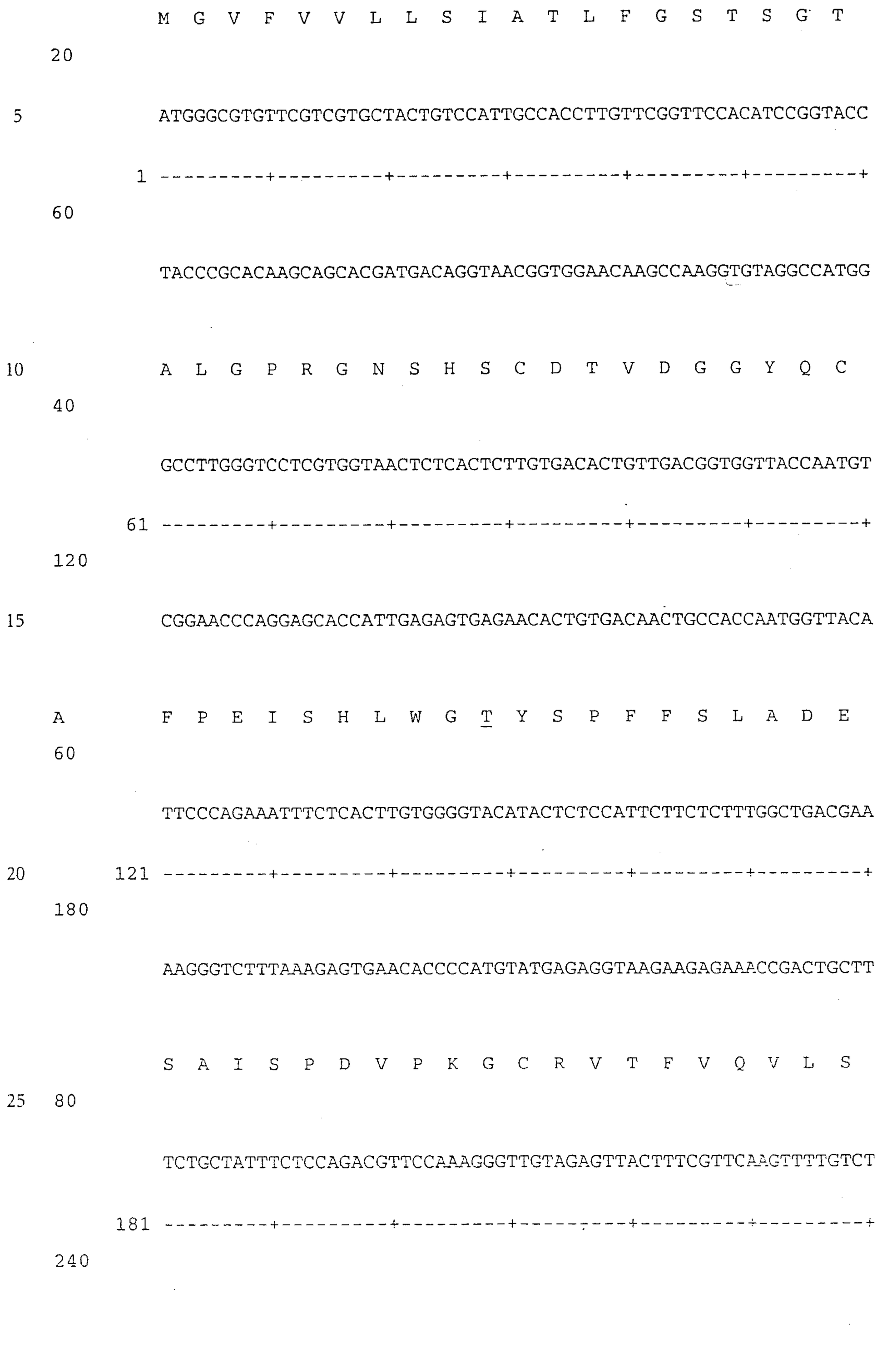
N W A E C F A * 467 (SEQ ID No. 105)

AACTGGGCTGAATGTTTCGCTTAA (SEQ ID No. 106)

1381 ---+---------+---------+ 1410

TTGACCCGACTTACAAAGCGAATT

Figure 9


M G V F V V L L S I A T L F G S T S G T
20
5 ATGGGCGTGTTCGTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCCACATCCGGTACC
1 --------+---------+---------+---------+---------+---------+
60
TACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGTGTAGGCCATGG
10 A L G P R G N S H S C D T V D G G Y Q C
40
GCCTTGGGTCCTCGTGGTAACTCTCACTCTTGTGACACTGTTGACGGTGGTTACCAATGT
61 --------+---------+---------+---------+---------+---------+
120
15 CGGAACCCAGGAGCACCATTGAGAGTGAGAACACTGTGACAACTGCCACCAATGGTTACA
A F P E I S H L W G T Y S P F F S L A D E
60
TTCCCAGAAATTTCTCACTTGTGGGGTACATACTCTCCATTCTTCTCTTTGGCTGACGAA
20 121 --------+---------+---------+---------+---------+---------+
180
AAGGGTCTTTAAAGAGTGAACACCCCATGTATGAGAGGTAAGAAGAGAAACCGACTGCTT
S A I S P D V P K G C R V T F V Q V L S
25 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

5               AGACACGGTGCTAGATACCCAACTTCTTCTGCGTCTAAGGCGTACTCTGCTTTGATTGAA
          241 --------+--------+--------+--------+--------+--------+--------+
    300

                TCTGTGCCACGATCTATGGGTTGAAGAAGACGCAGATTCCGCATGAGACGAAACTAACTT

10              A  I  Q  K  N  A  T  A  F  K  G  K  Y  A  F  L  K  T  Y  N
    120

                GCTATTCAAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGAAGACTTACAAC
          301 --------+--------+--------+--------+--------+--------+--------+
    360

15              CGATAAGTTTTCTTGCGATGACGAAAGTTCCCATTCATGCGAAAGAACTTCTGAATGTTG

    A           Y  T  L  G  A  D  D  L  T  P  F  G  E  Q  Q  M  V  N  S  G
    140

                TACACTTTGGGTGCTGACGACTTGACTCCATTCGGTGAACAACAAATGGTTAACTCTGGT

20        361 --------+--------+--------+--------+--------+--------+--------+
    420

                ATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTGTTGTTTACCAATTGAGACCA

                I  K  F  Y  R  R  Y  K  A  L  A  R  K  I  V  P  F  I  R  A
25  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

5               TCTGGTTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTTTCCAATCTGCT

        481 --------+---------+---------+---------+---------+---------+---------+
    540

                AGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAAAGGTTAGACGA

10              K  L  A  D  P  G  A  N  P  H  Q  A  S  P  V  I  N  V  I  I
    200

                AAGTTGGCTGACCCAGGTGCTAACCCACACCAAGCTTCTCCAGTTATTAACGTTATTATT

        541 --------+---------+---------+---------+---------+---------+---------+
    600

15              TTCAACCGACTGGGTCCACGATTGGGTGTGGTTCGAAGAGGTCAATAATTGCAATAATAA

                P  E  G  A  G  Y  N  N  T  L  D  H  G  L  C  T  A  F  E  E
    220

                CCAGAAGGTGCTGGTTACAACAACACTTTGGACCACGGTTTGTGTACTGCTTTCGAAGAA

20          601 --------+---------+---------+---------+---------+---------+---------+
    660

                GGTCTTCCACGACCAATGTTGTTGTGAAACCTGGTGCCAAACACATGACGAAAGCTTCTT

                S  E  L  G  D  D  V  E  A  N  F  T  A  V  F  A  P  P  I  R
25  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

5               GCTAGATTGGAAGCTCACTTGCCAGGTGTTAACTTGACTGACGAAGACGTTGTTAACTTG
        721 --------+--------+--------+--------+--------+--------+--------+
    780

                CGATCTAACCTTCGAGTGAACGGTCCACAATTGAACTGACTGCTTCTGCAACAATTGAAC

10              M  D  M  C  P  F  D  T  V  A  R  T  S  D  A  T  Q  L  S  P
    280

                ATGGACATGTGTCCATTCGACACTGTTGCTAGAACTTCTGACGCTACTCAATTGTCTCCA
        781 --------+--------+--------+--------+--------+--------+--------+
    840

15              TACCTGTACACAGGTAAGCTGTGACAACGATCTTGAAGACTGCGATGAGTTAACAGAGGT

                F  C  D  L  F  T  H  D  E  W  I  Q  Y  D  Y  L  Q  S  L  G
    300

                TTCTGTGACTTGTTCACTCACGACGAATGGATTCAATACGACTACTTGCAATCTTTGGGT
20      841 --------+--------+--------+--------+--------+--------+--------+
    900

                AAGACACTGAACAAGTGAGTGCTGCTTACCTAAGTTATGCTGATGAACGTTAGAAACCCA

                K  Y  Y  G  Y  G  A  G  N  P  L  G  P  A  Q  G  V  G  F  V
25  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

5               AACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAAGACCACACTTCTACTAACCAC

         961 --------+--------+--------+--------+--------+--------+--------+
    1020

                TTGCTTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAAGATGATTGGTG

10              T  L  D  S  N  P  A  T  F  P  L  N  A  T  L  Y  A  D  F  S
    360

                ACTTTGGACTCTAACCCAGCTACTTTCCCATTGAACGCTACTTTGTACGCTGACTTCTCT

        1021 --------+--------+--------+--------+--------+--------+--------+
    1080

15              TGAAACCTGAGATTGGGTCGATGAAAGGGTAACTTGCGATGAAACATGCGACTGAAGAGA

                H  D  N  T  M  V  S  I  F  F  A  L  G  L  Y  N  G  T  K  P
    380

                CACGACAACACTATGGTTTCTATTTTCTTCGCTTTGGGTTTGTACAACGGTACTAAGCCA

20      1081 --------+--------+--------+--------+--------+--------+--------+
    1140

                GTGCTGTTGTGATACCAAAGATAAAAGAAGCGAAACCCAAACATGTTGCCATGATTCGGT

                L  S  T  T  S  V  E  S  I  E  E  T  D  G  Y  S  A  S  W  T
25  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

5               GTTCCATTCGCTGCTAGAGCTTACGTTGAAATGATGCAATGTGAAGCTGAAAAGGAACCA

        1201 --------+---------+---------+---------+---------+---------+
    1260

                CAAGGTAAGCGACGATCTCGAATGCAACTTTACTACGTTACACTTCGACTTTTCCTTGGT

10              L  V  R  V  L  V  N  D  R  V  V  P  L  H  G  C  G  V  D  K
    440

                TTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTGGTGTTGACAAG

        1261 --------+---------+---------+---------+---------+---------+
    1320

15              AACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACACCACAACTGTTC

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

                TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTAGATCTGGTGGT

20          1321 --------+---------+---------+---------+---------+---------+
    1380

                AACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGATCTAGACCACCA

                N  W  E  E  C  F  A  *     467     (SEQ ID No. 107)

25              AACTGGGAAGAATGTTTCGCTTAA          (SEQ ID No. 108)

            1381 --------+---------+---- 1404

                TTGACCCTTCTTACAAAGCGAATT
```

Figure 10

```
                M  G  V  F  V  V  L  L  S  I  A  T  L  F  G  S  T  S  G  T   20

                ATGGGGGTTTTCGTCGTTCTATTATCTATCGCGACTCTGTTCGGCAGCACATCGGGCACT

5          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

10        61 ---------+---------+---------+---------+---------+---------+ 120

                CGCGACCCGGGGGCACCTTTAGTGAGGTTCAGGACGCTATGCCATCTGGATCCCATCGTC

                C  S  P  A  T  S  H  L  W  G  T  Y  S  P  Y  F  S  L  E  D   60

                TGCTCCCCTGCGACTTCTCATCTATGGGGCACGTACTCGCCATaCTTTTCGCTCGAGGAC

15       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

20       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

25       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

5 301 ---------+---------+---------+---------+---------+---------+ 360

TGCCGCTAGGTCCGGTTACGGTGGCTGAAGTTCCCGTTCAtgCGGAAAAACTTCTGCATG

N Y T L G A D D L T P F G E Q Q L V N S 140

AACTATACTCTGGGTGCGGATGACCTCACTCCCTTTGGGGAGCAGCAGCTGGTGAACTCG

10 361 ---------+---------+---------+---------+---------+---------+ 420

TTGATATGAGACCCACGCCTACTGGAGTGAGGGAAACCCCTCGTCGTCGACCACTTGAGC

G I K F Y Q R Y K A L A R S V V P F I R 160

GGCATCAAGTTCTACCAGAGGTACAAGGCTCTGGCGCGCAGTGTGGTGCCGTTTATTCGC

15 421 ---------+---------+---------+---------+---------+---------+ 480

CCGTAGTTCAAGATGGTCTCCATGTTCCGAGACCGCGCGTCACACCACGGCAAATAAGCG

A S G S D R V I A S G E K F I E G F Q Q 180

GCCTCAGGCTCGGACCGGGTTATTGCTTCGGGAGAGAAGTTCATCGAGGGGTTCCAGCAG

20 481 ---------+---------+---------+---------+---------+---------+ 540

CGGAGTCCGAGCCTGGCCCAATAACGAAGCCCTCTCTTCAAGTAGCTCCCCAAGGTCGTC

A K L A D P G A T N R A A P A I S V I I 200

GCGAAGCTGGCTGATCCTGGCGCGACGAACCGCGCCGCTCCGGCGATTAGTGTGATTATT

25 541 ---------+---------+---------+---------+---------+---------+ 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

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

5          901 ---------+---------+---------+---------+---------+---------+ 960

                TTCATGATGCCGATGCCGCGTCCGTTGGGAGACCCTGGCCGAGTCCCCTATCCCAAGTGG

                   N  E  L  I  A  R  L  T  R  S  P  V  Q  D  H  T  S  T  N  S
    340

10              AACGAGCTGATTGCCCGGTTGACgCGTTCGCCAGTGCAGGACCACACCAGCACTAACTCG

           961 ---------+---------+---------+---------+---------+---------+
    1020

                TTGCTCGACTAACGGGCCAACTGcGCAAGCGGTCACGTCCTGGTGTGGTCGTGATTGAGC

15                 T  L  V  S  N  P  A  T  F  P  L  N  A  T  M  Y  V  D  F  S
    360

                ACTCTAGTCTCCAACCCGGCCACCTTCCCGTTGAACGCTACCATGTACGTCGACTTTTCA

           1021 ---------+---------+---------+---------+---------+---------+
    1080

20              TGAGATCAGAGGTTGGGCCGGTGGAAGGGCAACTTGCGATGGTACATGCAGCTGAAAAGT

                   H  D  N  S  M  V  S  I  F  F  A  L  G  L  Y  N  G  T  E  P
    380

                CACGACAACAGCATGGTTTCCATCTTCTTTGCATTGGGCCTGTACAACGGCACTGAACCC

25         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

5       1141 ---------+---------+---------+---------+---------+---------+
    1200

                AACAGGGCCTGGAGCCACCTTTCGCGGTTCCTTAACCTACCCATAAGACGTAGGACCCAC

                V  P  F  G  A  R  A  Y  F  E  T  M  Q  C  K  S  E  K  E  P
10  420

                GTGCCTTTCGGCGCGCGAGCCTACTTCGAGACGATGCAATGCAAGTCGGAAAAGGAGCCT

        1201 ---------+---------+---------+---------+---------+---------+
    1260

                CACGGAAAGCCGCGCGCTCGGATGAAGCTCTGCTACGTTACGTTCAGCCTTTTCCTCGGA

15

                L  V  R  A  L  I  N  D  R  V  V  P  L  H  G  C  D  V  D  K
    440

                CTTGTTCGCGCTTTGATTAATGACCGGGTTGTGCCACTGCATGGCTGCGATGTGGACAAG

        1261 ---------+---------+---------+---------+---------+---------+
20  1320

                GAACAAGCGCGAAACTAATTACTGGCCCAACACGGTGACGTACCGACGCTACACCTGTTC

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

25              CTGGGGCGATGCAAGCTGAATGACTTTGTCAAGGGATTGAGTTGGGCCAGATCTGGGGGC

        1321 ---------+---------+---------+---------+---------+---------+
    1380

                GACCCCGCTACGTTCGACTTACTGAAACAGTTCCCTAACTCAACCCGGTCTAGACCCCCG
```

N W G E C F S * 467 (SEQ ID No. 109)

AACTGGGGAGAGTGCTTTAGTTGA (SEQ ID No. 110)

1381 ---------+---------+---- 1404

5 TTGACCCCTCTCACGAAATCAACT

Figure 11

CP-1

Eco RI M G V F V V L L S I A T L F G S T

10 **TATAT*GAATTC*ATGGGCGTGTTCGTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCCA**

1 ---------+---------+---------+---------+---------+---------+---------+ 60

ATATACTTAAGTACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGT

S G T A L G P R G N S H S C D T V D G G

15 **CATCCGGTACCGCCTTGGGTCCTC**GTGGTAATTCTCACTCTTGTGACACTGTTGACGGTG

61 ---------+---------+---------+---------+---------+---------+---------+ 120

GTAGG**CCATGGCGGAACCCAGGAGCACCATTAAGAGTGAGAACACTGTGACAACTGCCAC**

CP-2

CP-3

20 Y Q C F P E I S H L W G Q Y S P Y F S L

GTTACCAATGTTTCCCAGAAATTTCTCACTTGTGGGGTCAATACTCTCCATACTTCTCTT

121 ---------+---------+---------+---------+---------+---------+---------+ 180

CAATGGTTACAAAGGGTCTTTAAAGAGTGAACACCCCAGTTATGAGAGGTATGAAGAGAA

25 E D E S A I S P D V P D D C R V T F V Q

TGGAAGACGAATCTGCTATTTCTCCAGACGTTCCAGACGACTGTAGAGTTACTTTCGTTC

181 ---------+---------+---------+---------+---------+---------+---------+ 240

ACCTTCTGCTTAGAC**GATAAAGAGGTCTGCAAGGTCTGCTGACATCTCAATGAAAGCAAG**

CP-4.7

30 CP-5.7

V L S R H G A R Y P T D S K G K K Y S A

AAGTTTTGTCTAGACACGGT**GCTAGATACCCAACTgacTCTAAGggtAAGaagTACTCTG**

241 ---------+---------+---------+---------+---------+---------+ 300

TTCAAAACAGATCTGTGCCACGATCTATGGGTTGActgAGATTCccaTTCttcATGAGAC

5 L I E A I Q K N A T A F K G K Y A F L K

CTTTGATTGAAGCTATTCAAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGA

301 ---------+---------+---------+---------+---------+---------+ 360

GAAACTAACTTCGATAAGTTTTCTT**GCGATGACGAAAGTTCCCATTCATGCGAAAGAACT**

CP-6

10 CP-7

T Y N Y T L G A D D L T P F G E N Q M V

AGACTTACAACTACACTTTGGGTGCTGACG**ACTTGACTCCATTCGGTGAAAACCAAATGG**

361 ---------+---------+---------+---------+---------+---------+ 420

TCTGAATGTTGATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACC

15

N S G I K F Y R R Y K A L A R K I V P F

TTAACTCTGGTATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCAT

421 ---------+---------+---------+---------+---------+---------+ 480

AATTGAGACCATAATTCAAGATGTCTTCTATGTTC**CGAAACCGATCTTTCTAACAAGGTA**

20 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

TCATTAGAGCTTCTGGTTCTtctAGAGTTATTGCTTCTGC**TGAAAAGTTCATTGAAGGTT**

481 ---------+---------+---------+---------+---------+---------+ 540

25 **AGTAATCTCGAAGACCAAGAagaTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAA**

Q S A K L A D P G S Q P H Q A S P V I D

TCCAATCTGCTAAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTG

541 ---------+---------+---------+---------+---------+---------+ 600

30 AGGTTAGACGATTCAACCGACTGGGTCCAAGAGTTGGTGTGGTTC**GAAGAGGTCAATAAC**

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

ACGTTATTATTtctGAcgctTCTtctTACAACAACACTTTGGACccaGGT<u>ACTTGTACTG</u>

601 ---------+---------+---------+---------+---------+---------+ 660

TGCAATAATAAagaCTgcgaAGGagaATGTTGTTGTGAAACCTGggtCCATGAACATGAC

F E D S E L A D T V E A N F T A L F A P

CTTTCGAAGACTCTGAATTGgctGACactGTTGAAGCTAACTTCACTGCTTTGTTCGCTC

661 ---------+---------+---------+---------+---------+---------+---------+ 720

GAAAGCTTCTGAGACTTAACcgaCTGtgaCAACTTCGATTGAAGTGACGAAACAA**GCGAG**

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

TTAAC**AGAGGTAAGACACGAAACAAGTGAGTGCTGCTTACCTCTgtgATGCTGATGAACG**

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

AATCTTTGaag**AAGTACTACGGTcacGGTGCTGGTAACCCATTGGGTCCAactCAAGGTG**

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

AACCAAAGCGATTGC**TTAACTAACGATCTAACTGATCTAGAGGTCAAGTTCTGGTGTGAA**

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

5              GATGATTGGTGTGAAACCTGAGATTGGGTCGATGAAAGGGTAACTTGCGATGAAACATGC

                D  F  S  H  D  N  G  I  I  S  I  F  F  A  L  G  L  Y  N  G

               CTGACTTCTCTCACGACAACggtattATTTCTATTTTCTTCGCTTTGGGTTTGTACAACG

          1081 ---------+---------+---------+---------+---------+---------+
10   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

15             GTACTGCTCCATTGTCTACTACTTCTGTTGAATCTATTGAAGAAACTGACGGTTACTCTt

          1141 ---------+---------+---------+---------+---------+---------+
1200

               CATGACGAGGTAACAGATGATGAAGACAACTTAGATAACTTCTTTGACTGCCAATGAGAa

20              A  W  T  V  P  F  A  S  R  A  Y  V  E  M  M  Q  C  Q  A  E

               ctgctTGGACTGTTCCATTCgcttctAGAGCTTACGTTGAAATGATGCAATGTCAAGCTG
          1201 ---------+---------+---------+---------+---------+---------+
1260

               gacgaACCTGACAAGGTAAGcgaagaTCTCGAATGCAACTTTACTACGTTACAGTTCGAC

25                                                  CP-20
                                                         CP-21

                K  E  P  L  V  R  V  L  V  N  D  R  V  V  P  L  H  G  C  A

               AAAAGGAACCATTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTG
          1261 ---------+---------+---------+---------+---------+---------+
30   1320

                    TTTTCCTTGGTAACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACAC
```

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

CTGTTGACAAGTTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTA

1321 ---------+---------+---------+---------+---------+---------+ 1380

GACAACTGTTCAACCCATCTACATTCTCTCTGCTGAAGCAACTTC**CAAACAGAAAGCGAT**

S G G N W A E C F A * *Eco* RI CP-22 (SEQ ID No. 111)

GATCTGGTGGTAACTGGGCTGAATGTTTCGCT*TAA*GAATTCATATA (SEQ ID No. 112)

1381 ---------+---------+---------+---------+---------+------ 1426

CTAGACCACCATTGACCCGACTTACAAAGCGAATTCTTAAGTATAT

Figure 12

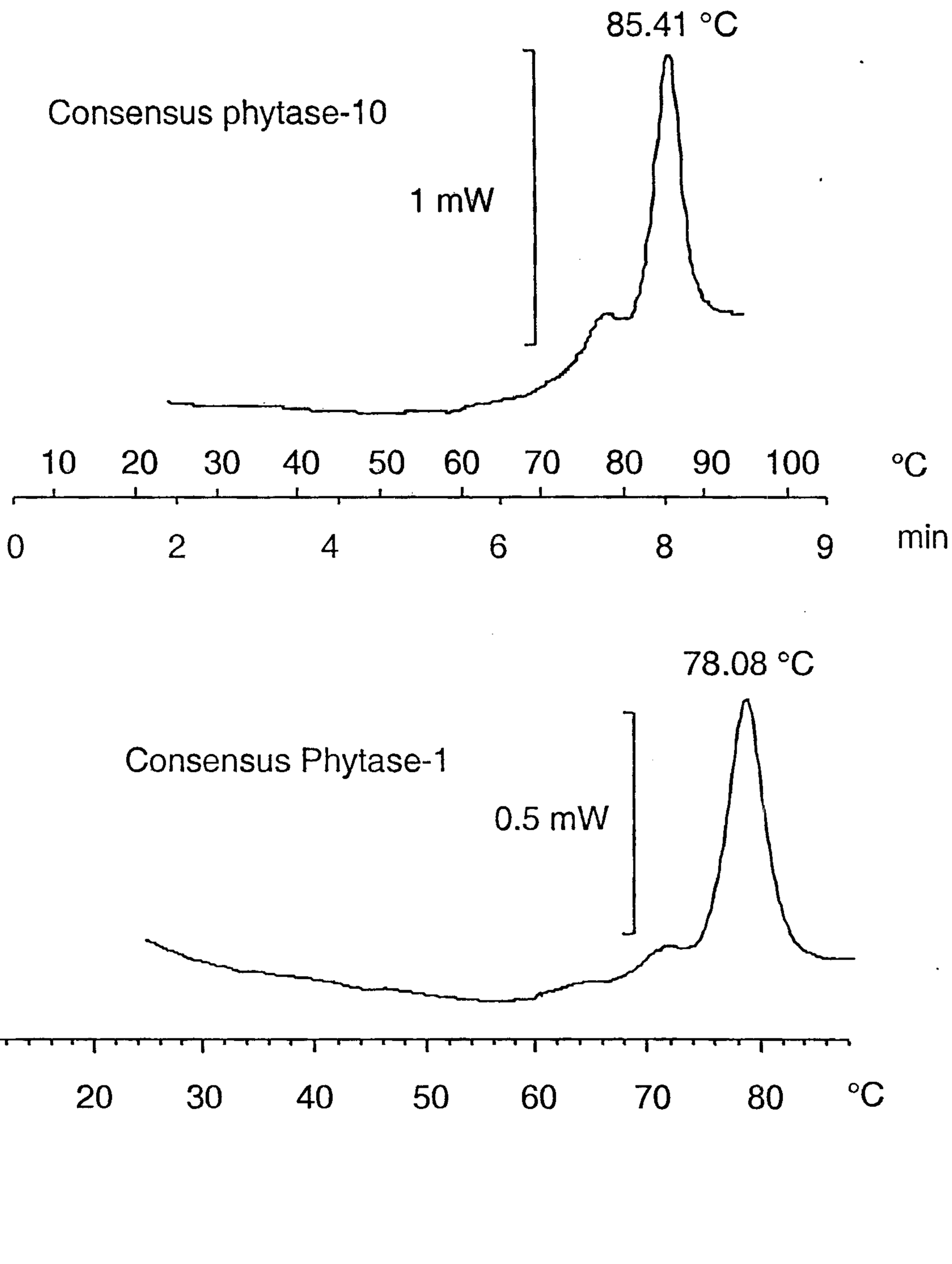
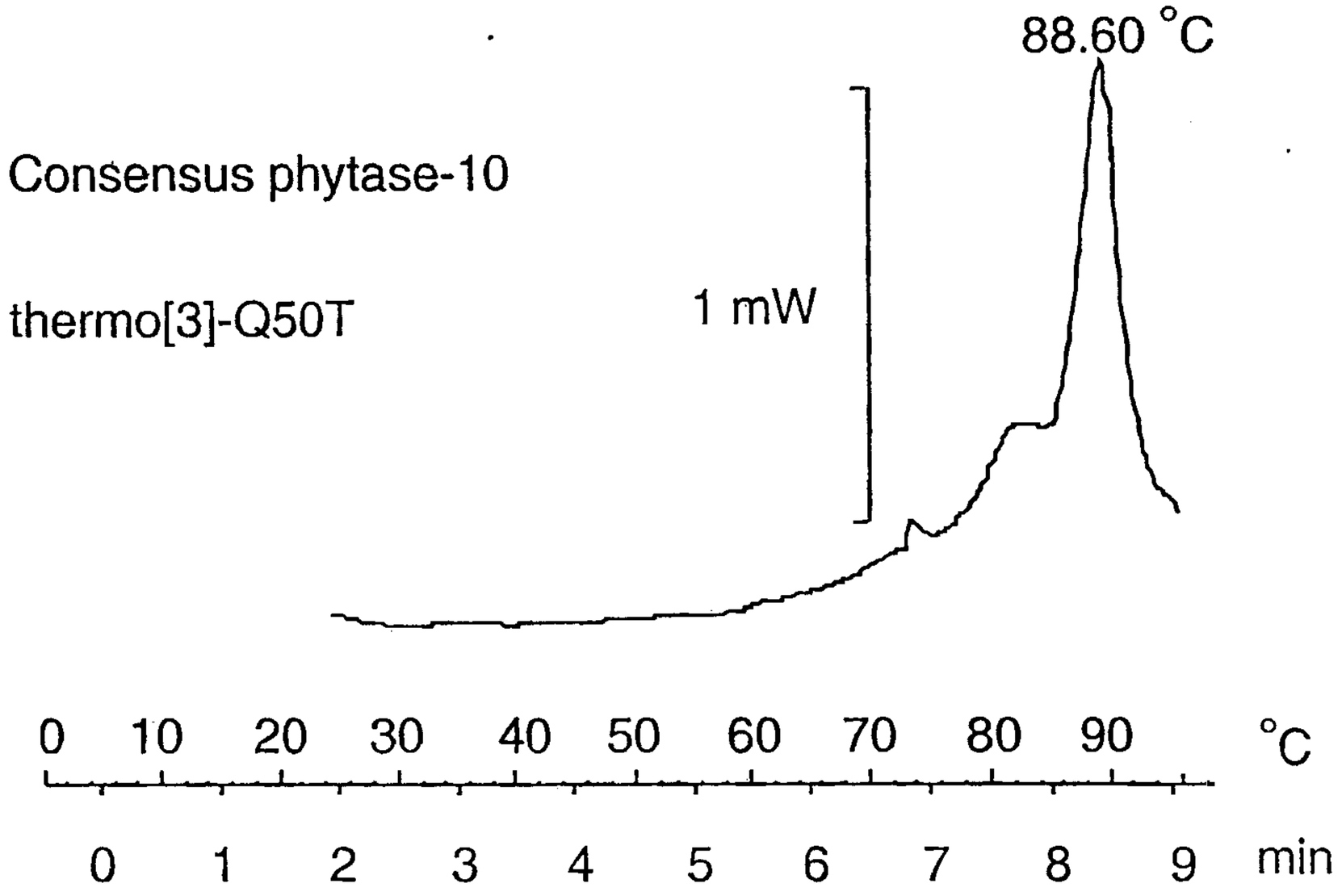

85.41 °C
Consensus phytase-10
1 mW
10 20 30 40 50 60 70 80 90 100 °C
0 2 4 6 8 9 min
78.08 °C
Consensus Phytase-1
0.5 mW
10 20 30 40 50 60 70 80 °C

Figure 13


88.60 °C
Consensus phytase-10
thermo[3]-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

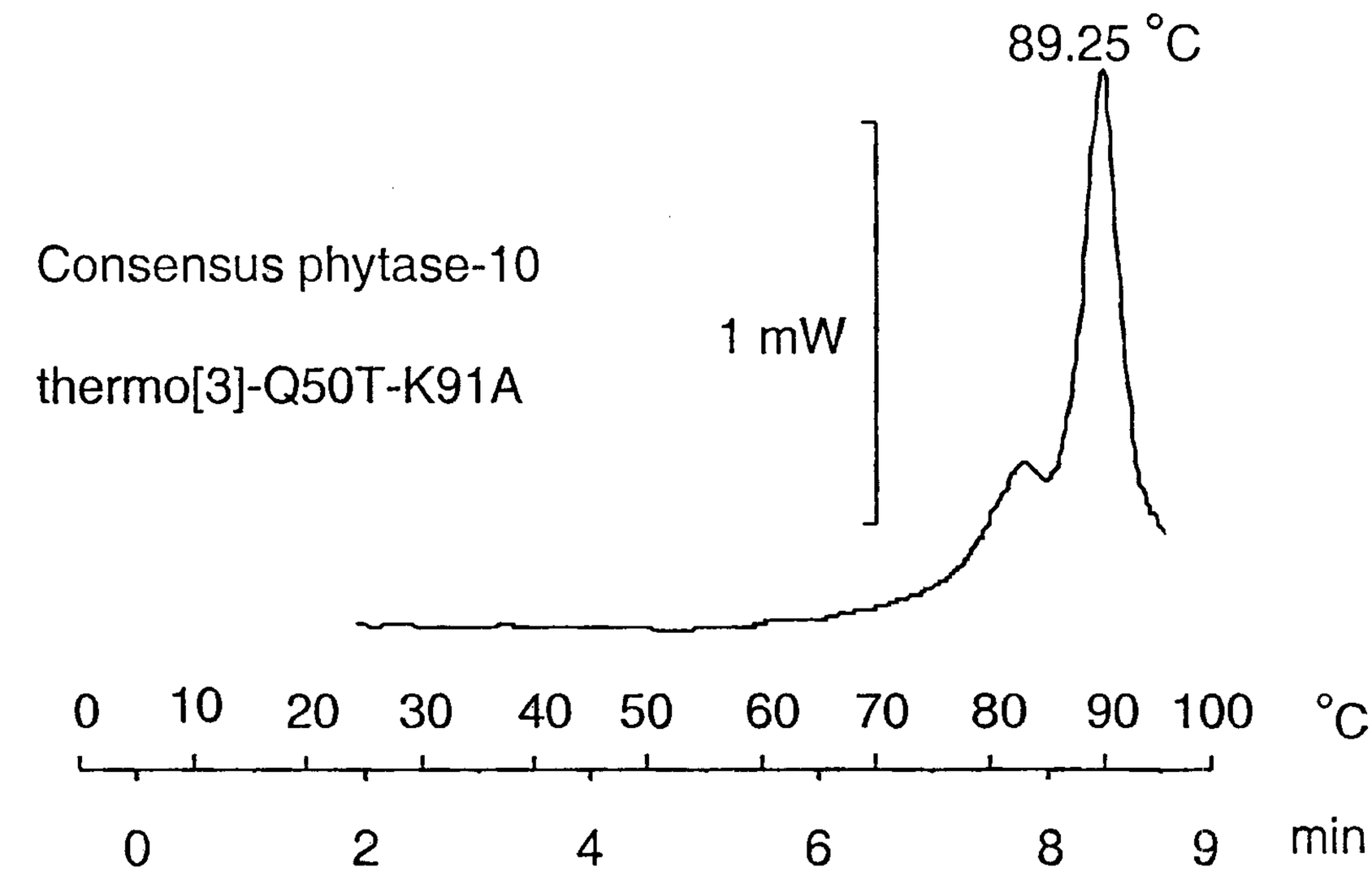

89.25 °C
Consensus phytase-10
thermo[3]-Q50T-K91A
1 mW
0 10 20 30 40 50 60 70 80 90 100 °C
0 2 4 6 8 9 min

Figure 14

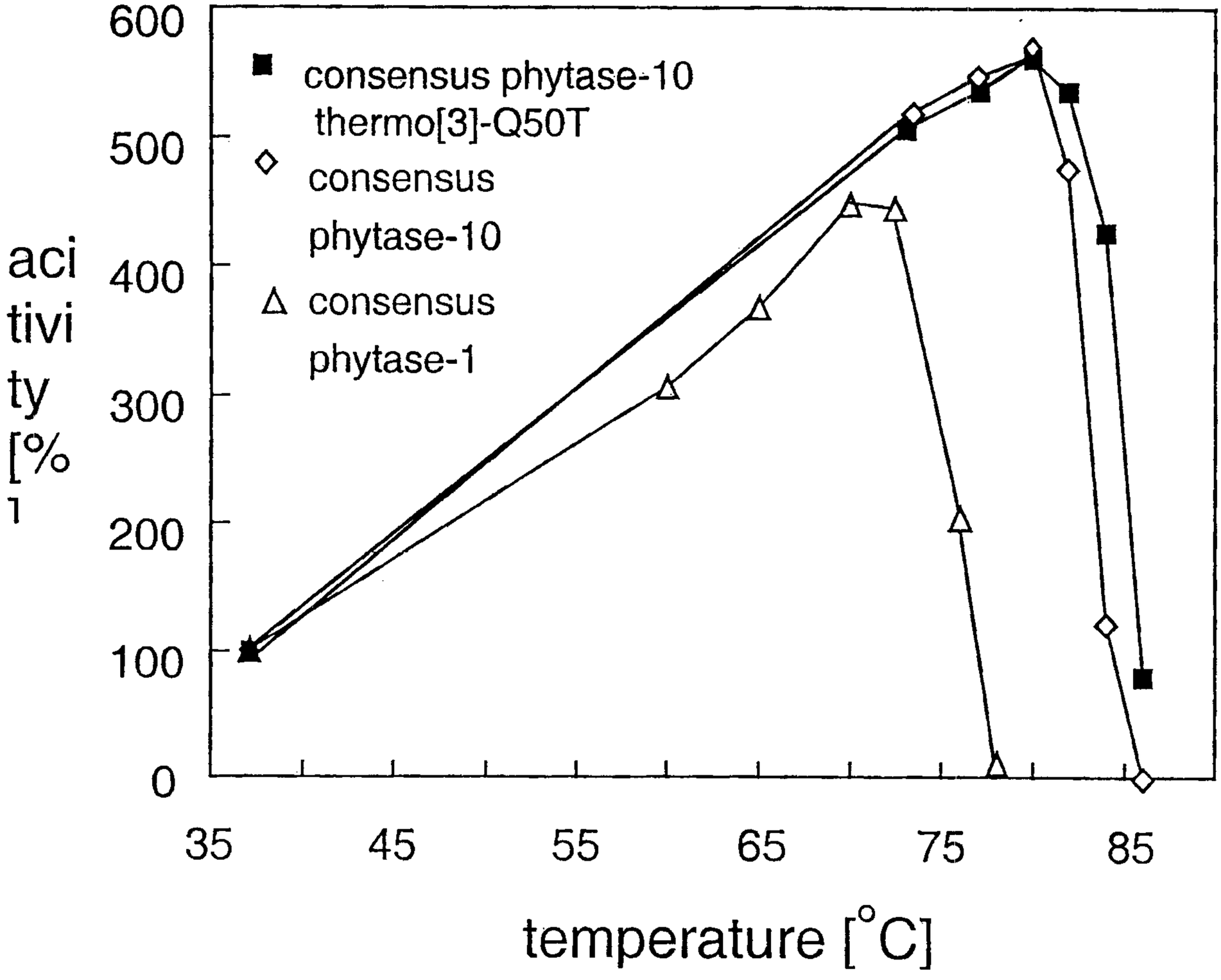
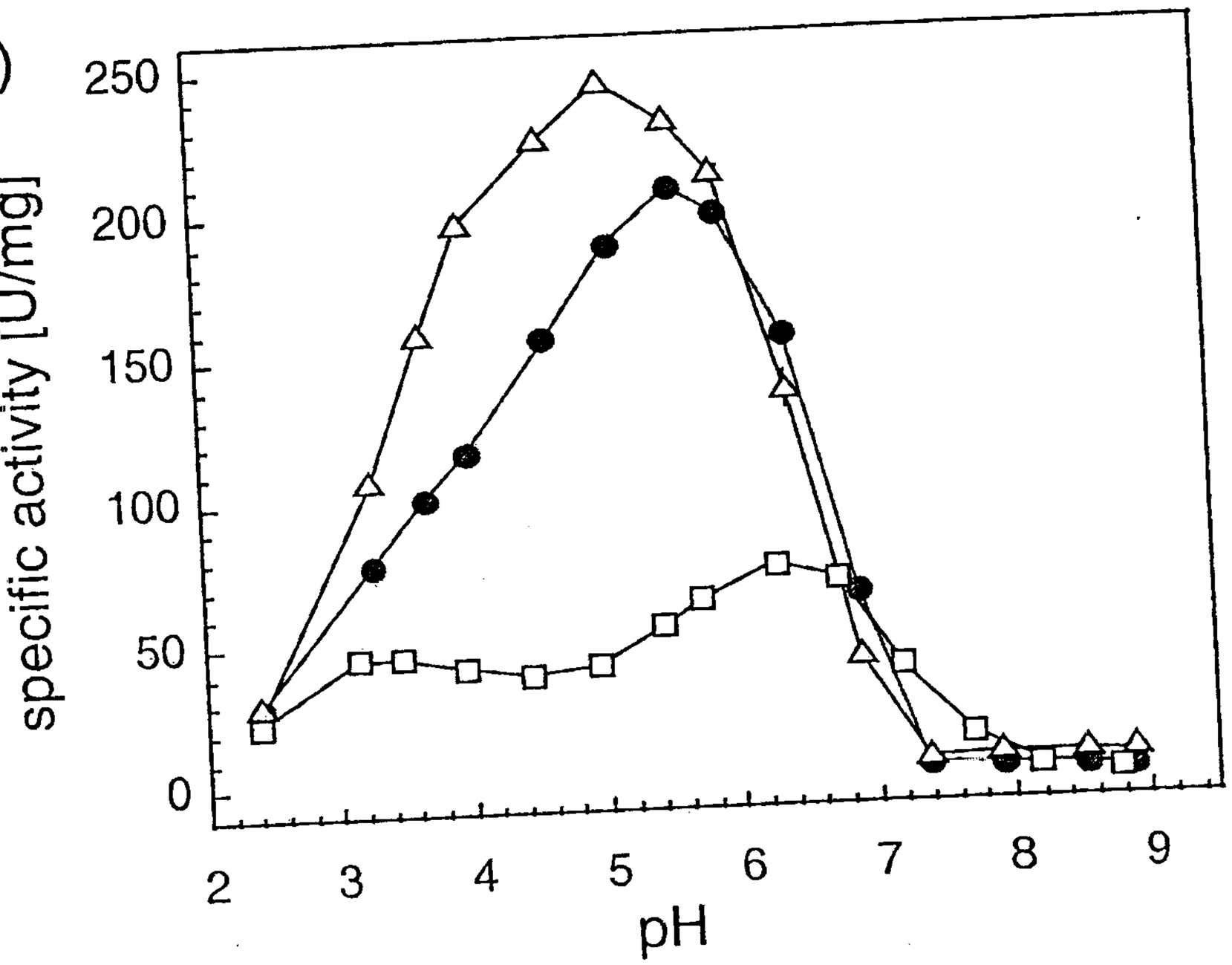

consensus phytase-10 thermo[3]-Q50T
consensus phytase-10
consensus phytase-1
activity [%]
600
500
400
300
200
100
0
35
45
55
65
75
85
temperature [°C]

Figure 15


a)
250
200
150
100
50
0
specific activity [U/mg]
2
3
4
5
6
7
8
9
pH

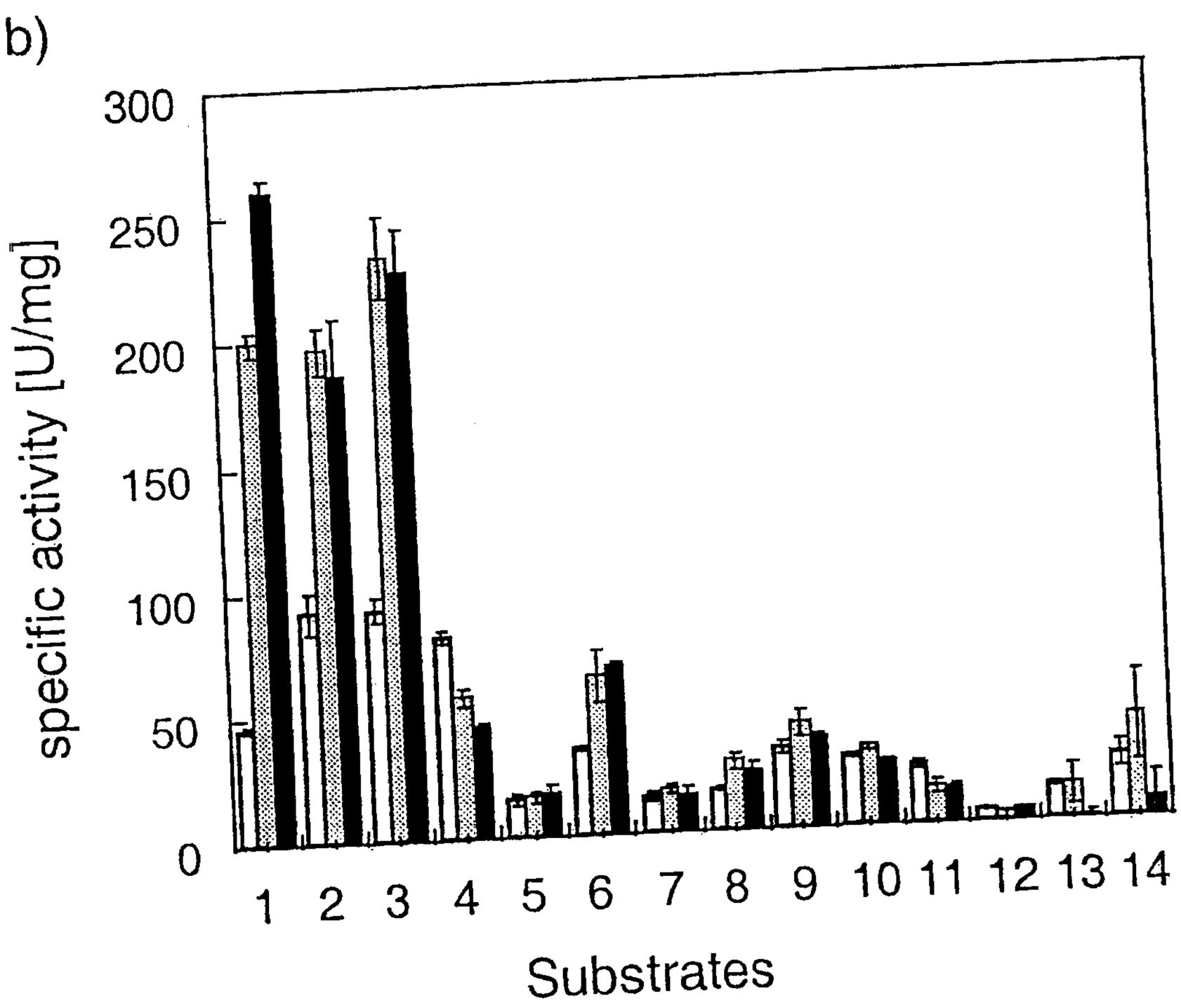
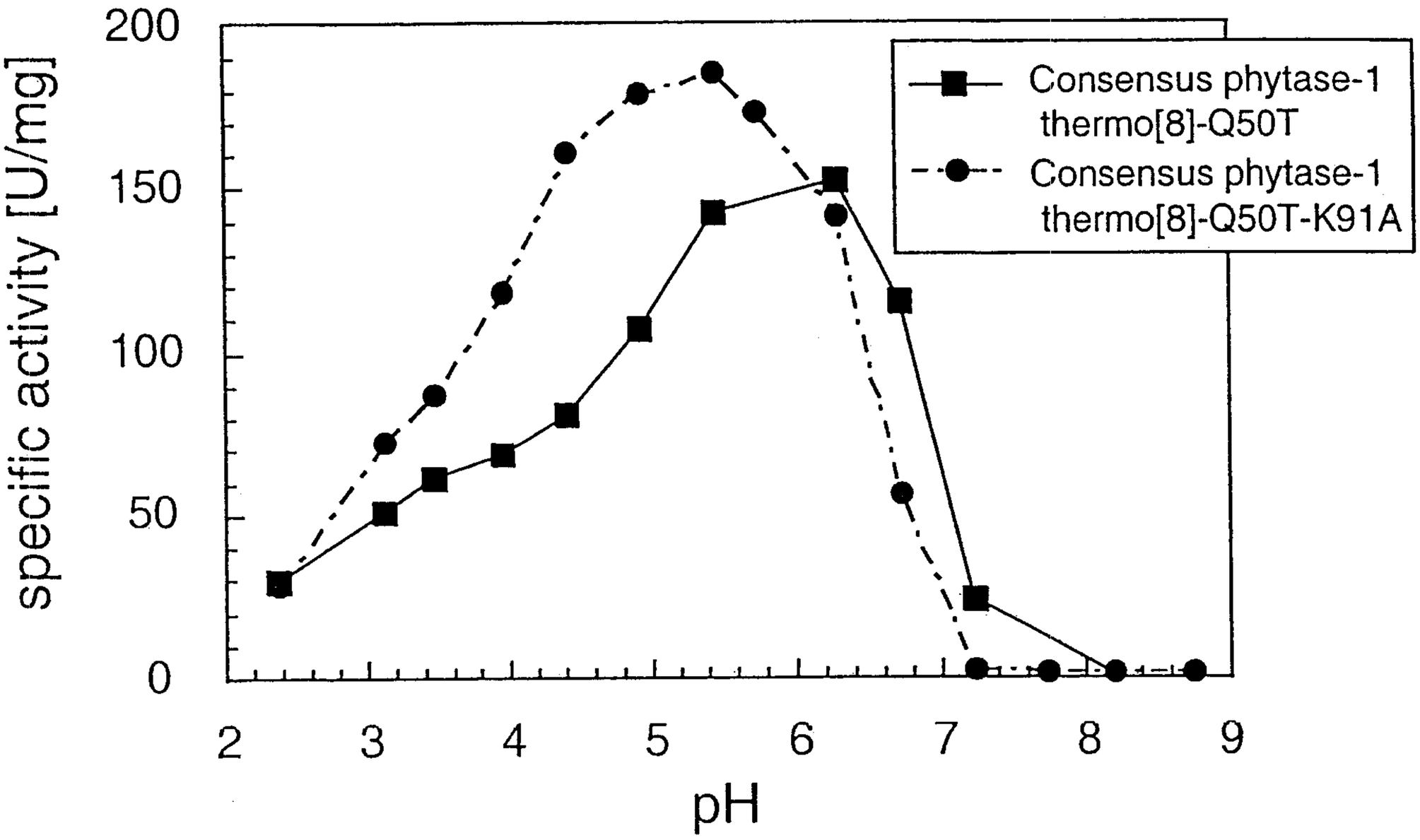

b)
300
250
200
150
100
50
0
specific activity [U/mg]
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Substrates

Figure 16


specific activity [U/mg]
200
150
100
50
0
2
3
4
5
6
7
8
9
pH
Consensus phytase-1 thermo[8]-Q50T
Consensus phytase-1 thermo[8]-Q50T-K91A

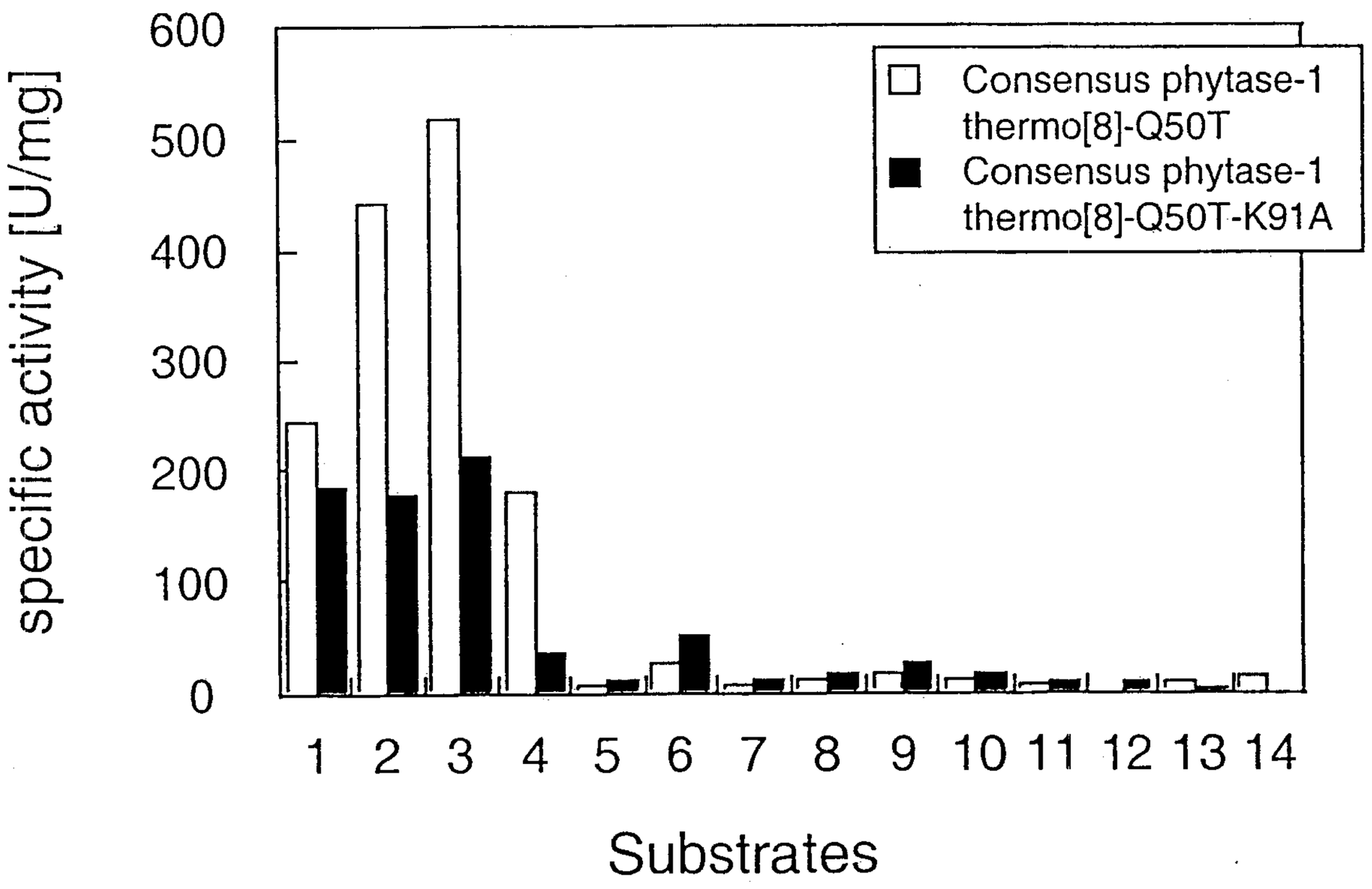

specific activity [U/mg]
600
500
400
300
200
100
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Substrates
Consensus phytase-1 thermo[8]-Q50T
Consensus phytase-1 thermo[8]-Q50T-K91A

Figure 17

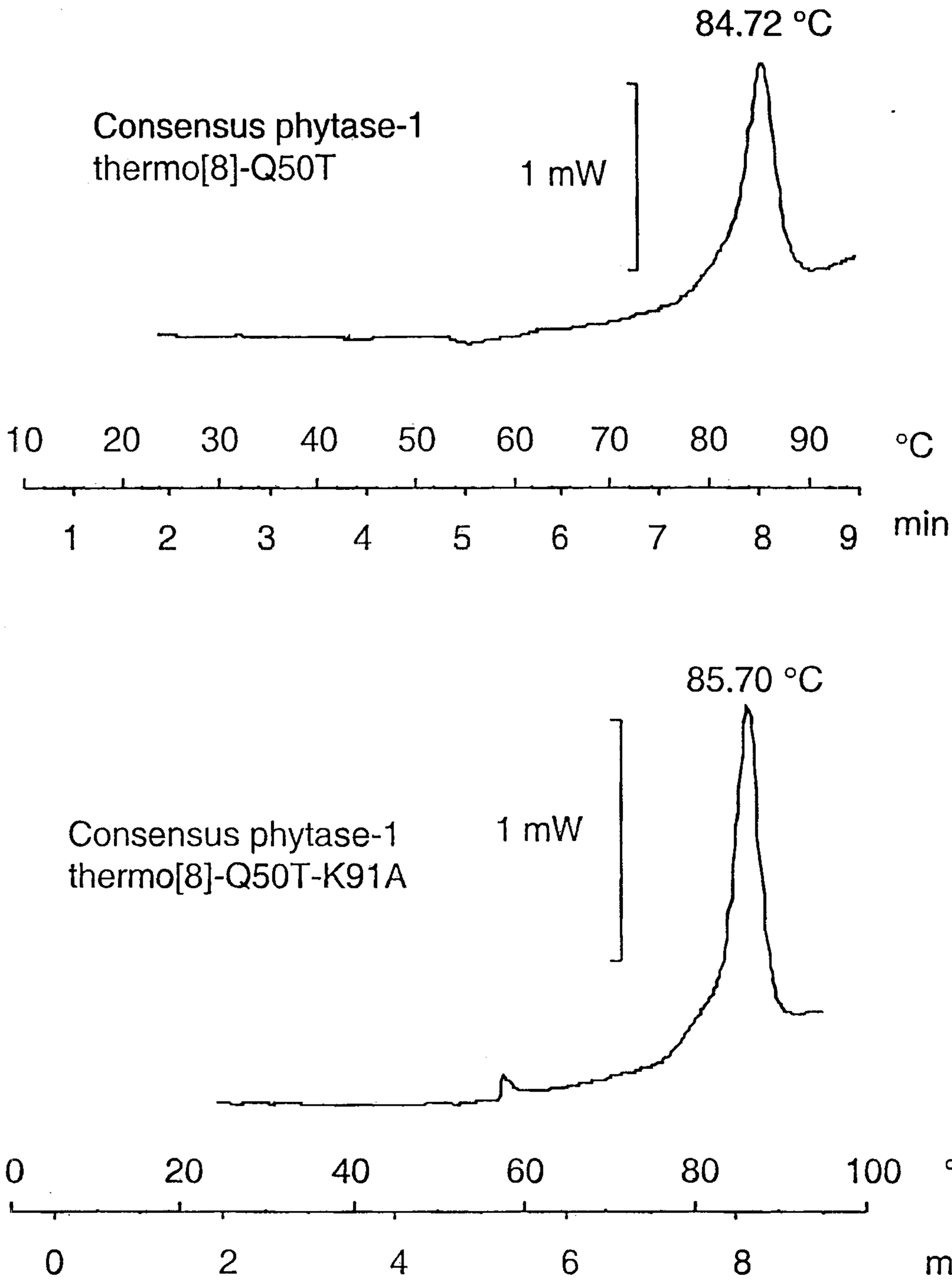

84.72 °C
Consensus phytase-1
thermo[8]-Q50T
1 mW
10 20 30 40 50 60 70 80 90 °C
1 2 3 4 5 6 7 8 9 min
85.70 °C
Consensus phytase-1
thermo[8]-Q50T-K91A
1 mW
0 20 40 60 80 100 °C
0 2 4 6 8 min

Figure 18

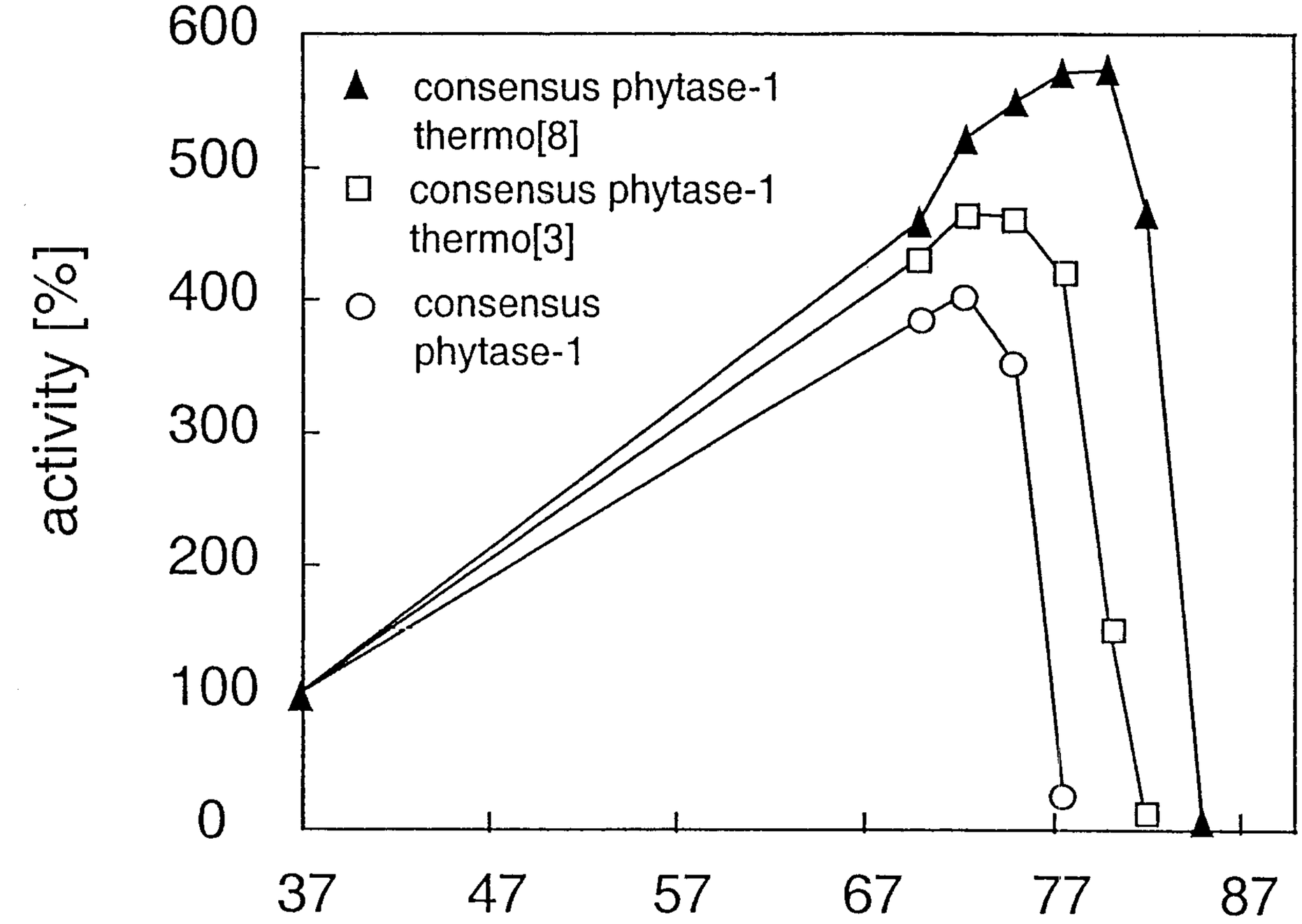
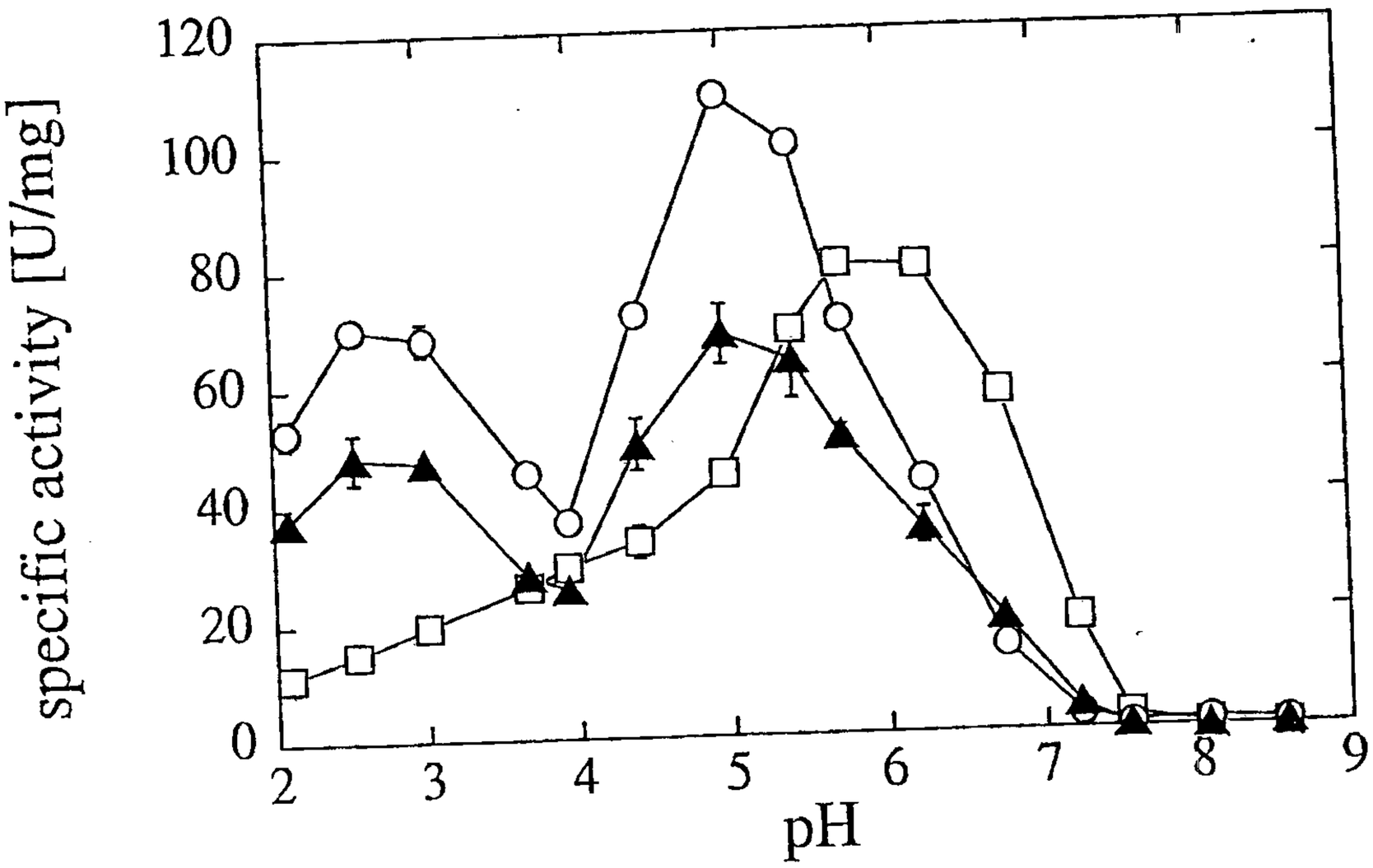

600
500
400
300
200
100
0
activity [%]
consensus phytase-1 thermo[8]
consensus phytase-1 thermo[3]
consensus phytase-1
37
47
57
67
77
87
temperature [°C]

Figure 19


specific activity [U/mg]
120
100
80
60
40
20
0
2
3
4
5
6
7
8
9
pH

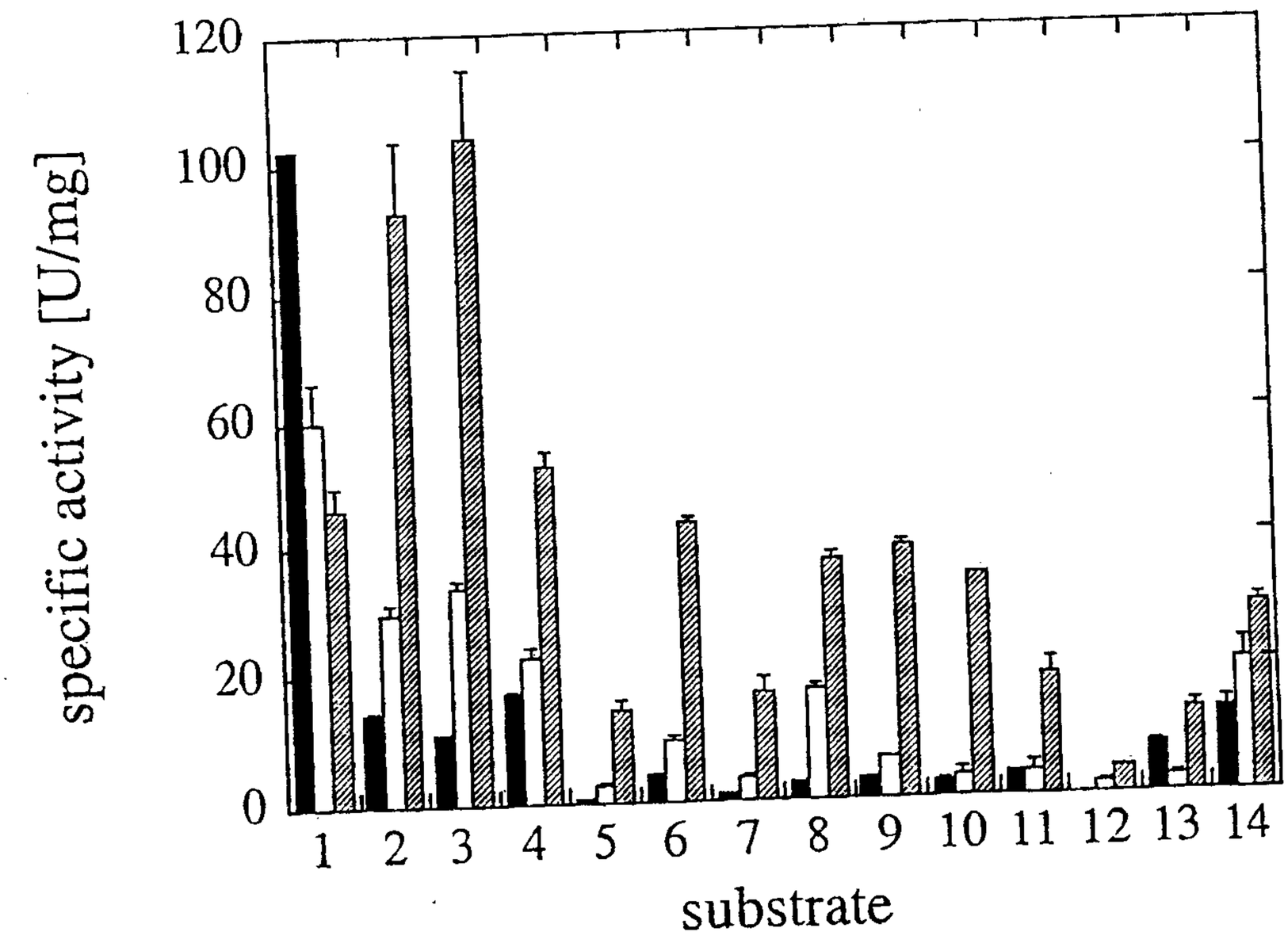
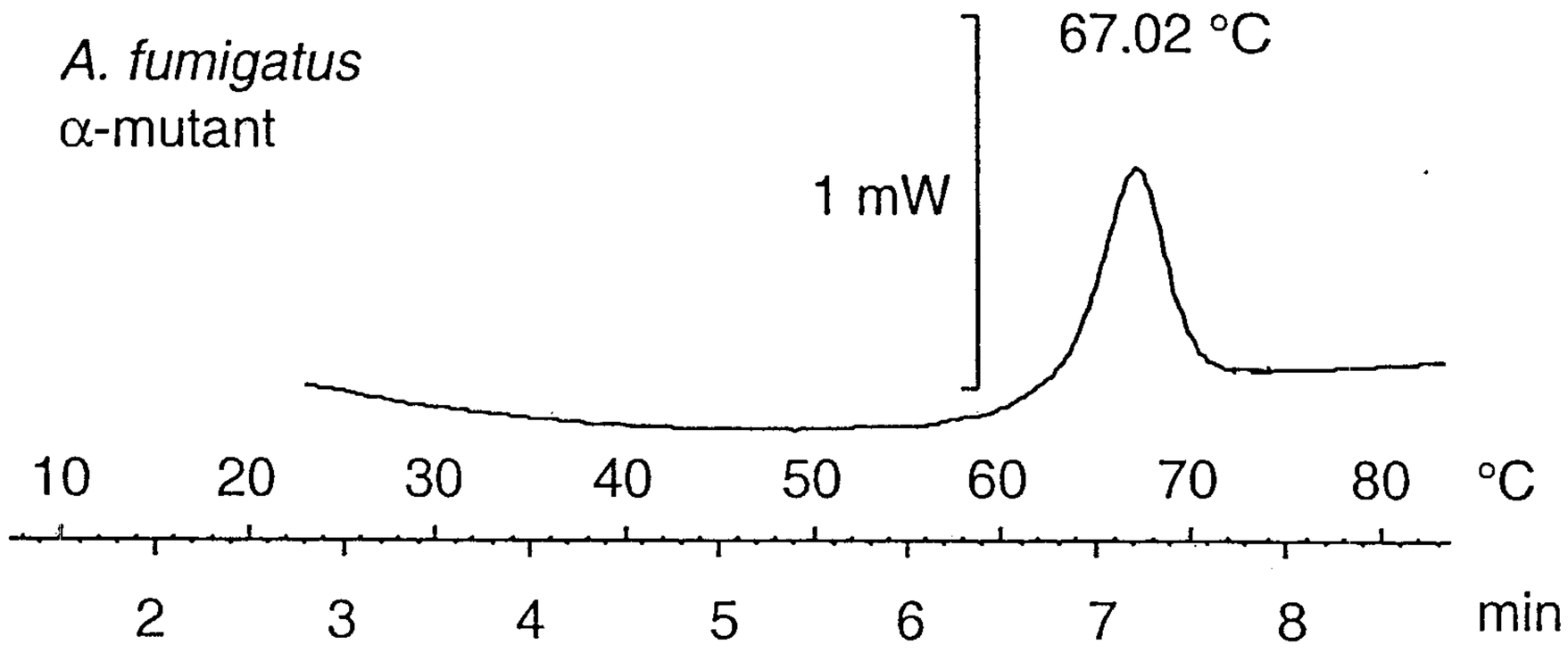

specific activity [U/mg]
120
100
80
60
40
20
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
substrate

Figure 20


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

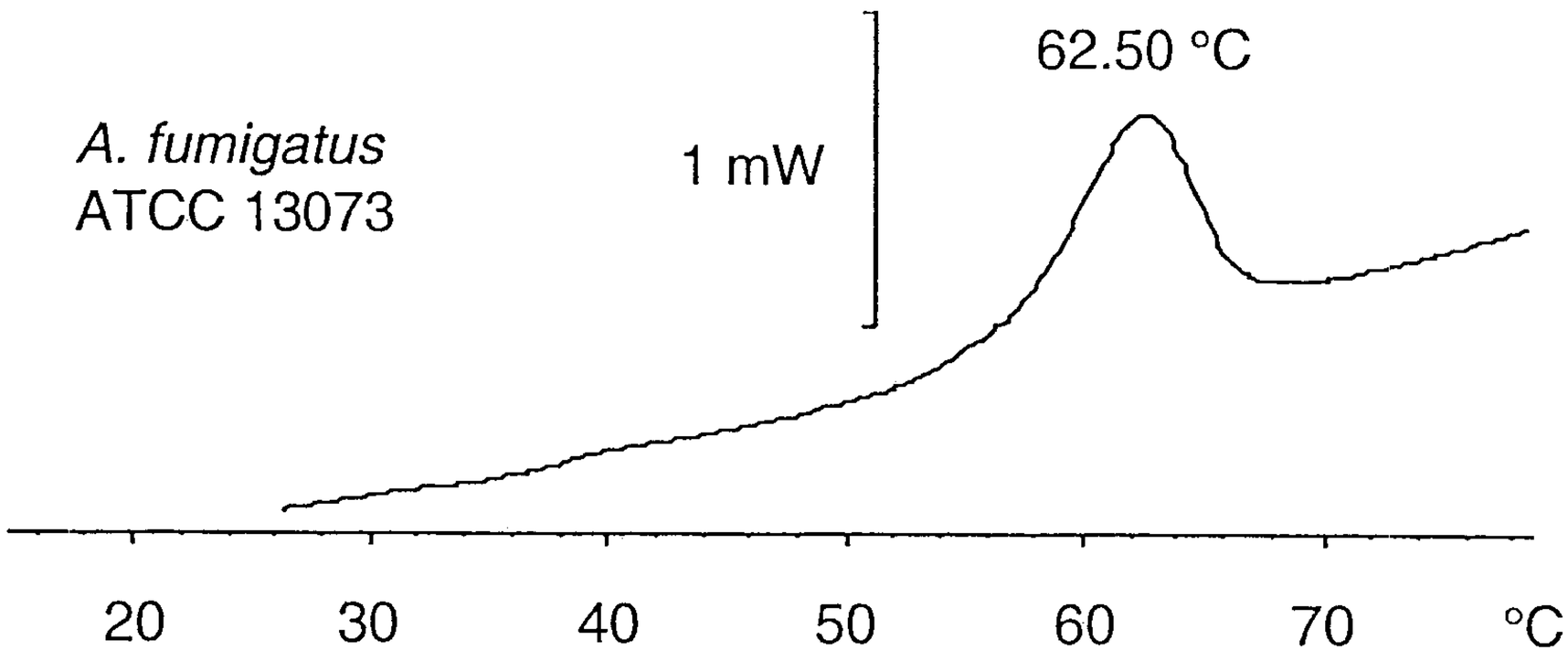

62.50 °C
A. fumigatus
ATCC 13073
1 mW
20
30
40
50
60
70
°C

Figure 21

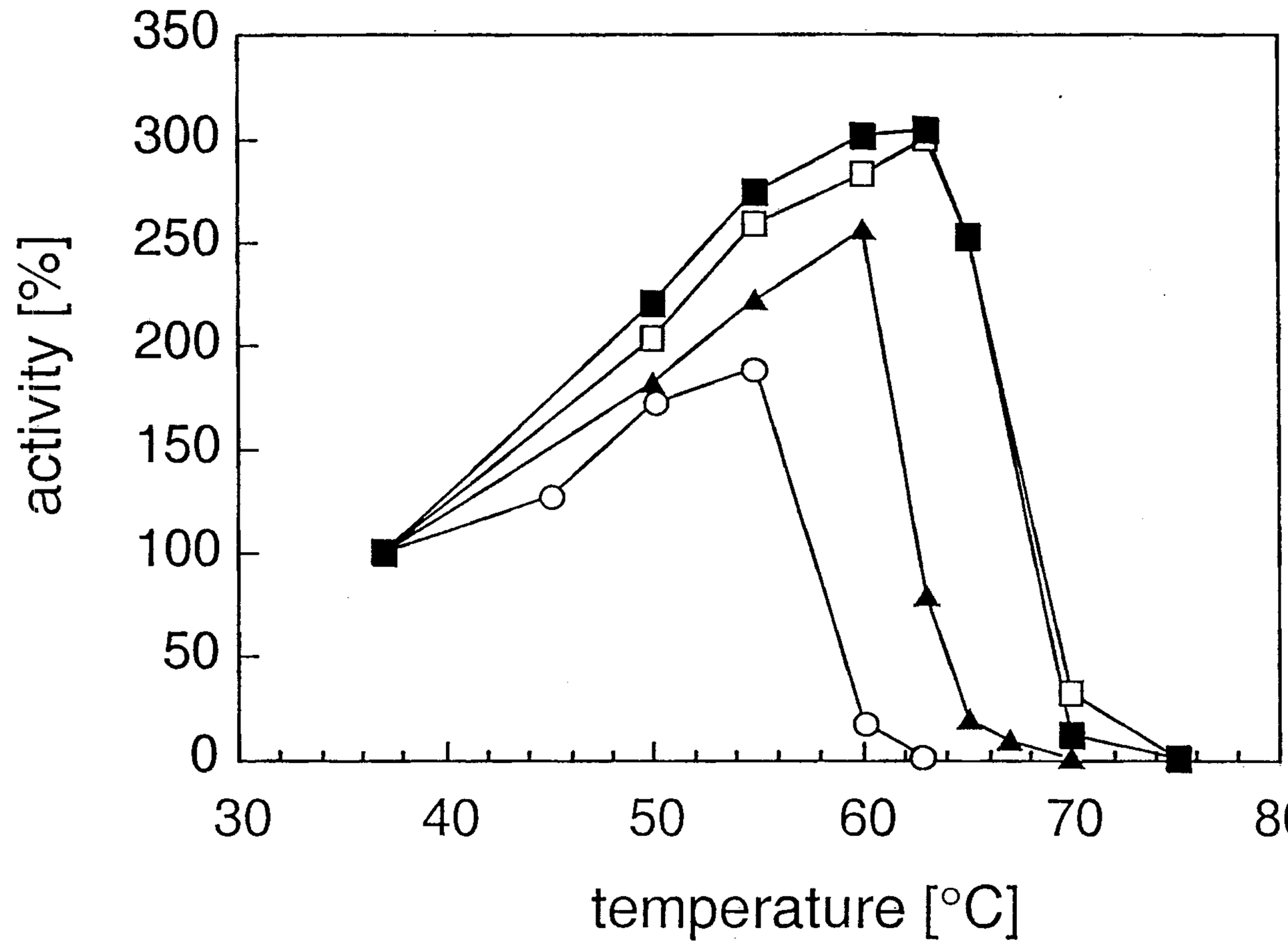

350
300
250
200
150
100
50
0
activity [%]
30
40
50
60
70
80
temperature [°C]

Figure 22

1 MGVFVVLLSI ATLFGSTSGT ALGPRGNSHS CDTVDGGYQC FPEISSNWSP

51 YSPYFSLADE SAISPDVPKG CRVTFVQVLQ RHGARFPTSG AATRISALIE

101 AIQKNATAFK GKYAFLKTYN YTLGADDLVP FGANQSSQAG IKFYRRYKAL

151 ARKIVPFIRA SGSDRVIDSA TNWIEGFQSA KLADPGANPH QASPVINVII

201 PEGAGYNNTL DHGLCTAFEE SELGDDVEAN FTAVFAPPIR ARLEAHLPGV

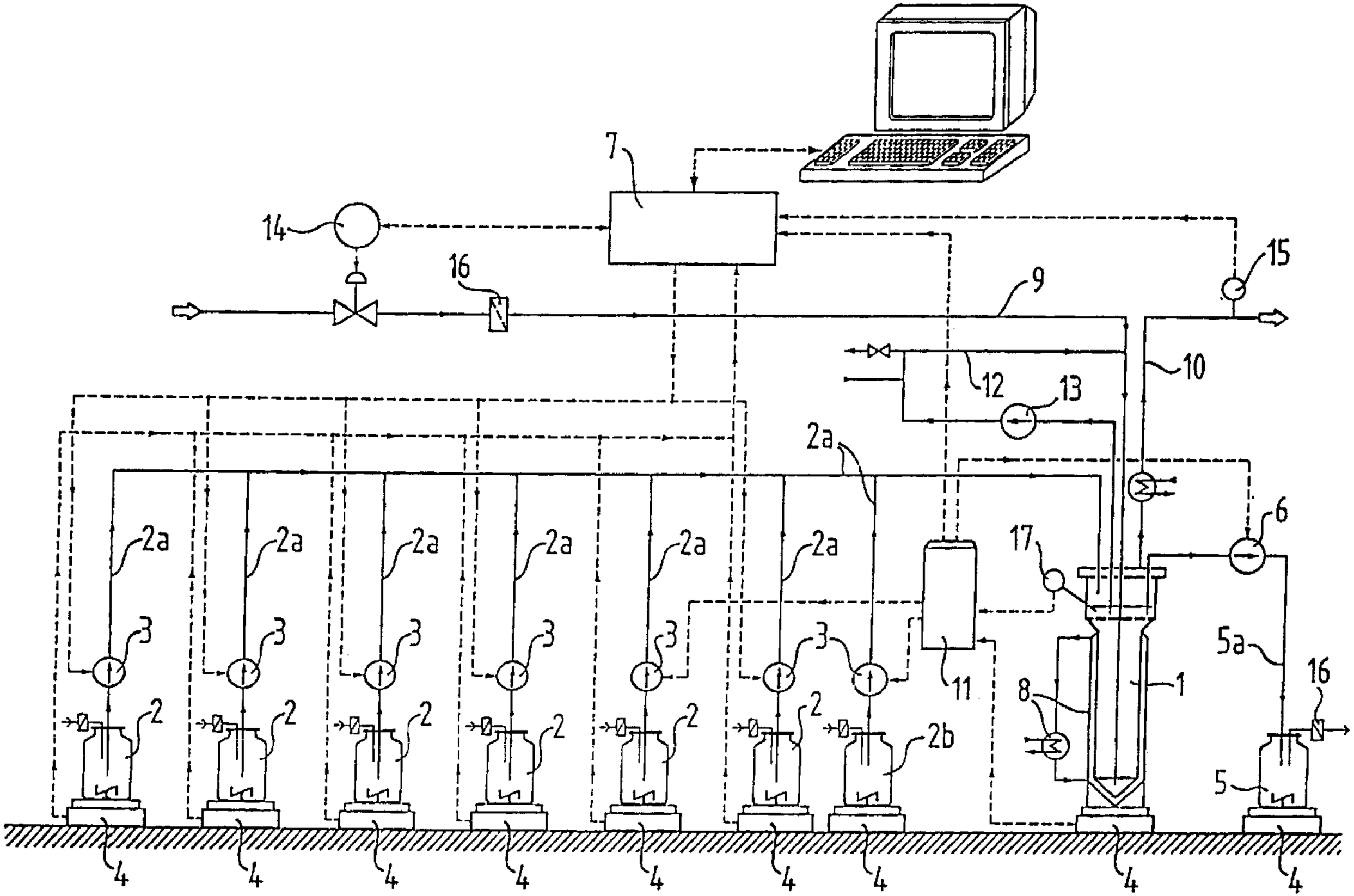
251 NLTDEDVVNL MDMCPFDTVA RTSDATELSP FCDLFTHDEW IQYDYLGDLD

301 KYYGTGAGNP LGPAQGVGFV NELIARLTHS PVQDHTSTNH TLDSNPATFP

351 LNATLYADFS HDNTMVAIFF ALGLYNGTKP LSTTSVESIE ETDGYSASWL

401 VPFSARMYVE MMQCEAEKEP LVRVLVNDRV VPLHGCGVDK LGRCKRDDFV

451 EGLSFARSGG NWEECFA (SEQ ID No. 113)


7
14
16
15
9
10
12
13
2a
6
17
3
2
2b
11
1
8
5a
5
16
4