

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
4 March 2004 (04.03.2004)

PCT

(10) International Publication Number
WO 2004/018662 A2

(51) International Patent Classification⁷: **C12N 9/42**,
9/24, A21D 8/04

(21) International Application Number:
PCT/EP2003/009147

(22) International Filing Date: 15 August 2003 (15.08.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
02102152.2 19 August 2002 (19.08.2002) EP
02102160.5 19 August 2002 (19.08.2002) EP
02102161.3 19 August 2002 (19.08.2002) EP
02102245.4 30 August 2002 (30.08.2002) EP
02102279.3 3 September 2002 (03.09.2002) EP
02102298.3 3 September 2002 (03.09.2002) EP

(71) Applicant (*for all designated States except US*): **DSM IP ASSETS B.V.** [NL/NL]; Het Overloon 1, NL-6411 TE HEERLEN (NL).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **FOLKERS, Ulrike** [DE/DE]; Treffauer Strasse 54, 81373 München (DE). **FRITZ, Andreas** [DE/DE]; Lochhamer Strasse 11, 82152 MARTINSRIED (DE). **GERHARD, Beatrix** [DE/DE]; Albrecht-Schweitzer 3, 85521 Ottobrunn (DE). **KLUG-BAUER, Sabine** [DE/DE]; Friedenspromenade 48 A, 81827 München (DE). **SPREAFICO, Fabio** [IT/DE]; Lochhamer Str. 11, 82152 MARTINSRIED (DE). **WAGNER, Christian** [DE/DE]; Lochhamer Str. 11, 82152

MARTINSRIED (DE). **BOER, DE, Lex** [NL/NL]; Mid-denweg 61, NL-2291 HV WATERINGEN (NL). **MEIMA, Roelf, Bernhard** [NL/NL]; Hoefslag 28, NL-3471 AH Kamerik (NL).

(74) Agent: **MATULEWICZ, Emil, Rudolf, Antonius**; DSM Intellectual Property, Delft Office (600-0240), P.O. Box 1, NL-2600 MA DELFT (NL).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL (HEMI)CELLULASES AND USES THEREOF

(57) Abstract: The invention relates to a newly identified polynucleotide sequence comprising a gene that encodes a novel (hemi)cellulase from *Aspergillus niger*. The invention features the full length nucleotide sequence of the novel gene, the cDNA sequence comprising the full length coding sequence of the novel (hemi)cellulase as well as the amino acid sequence of the full-length functional protein and functional equivalents thereof. The invention also relates to methods of using these enzymes in industrial processes and methods of diagnosing fungal infections. Also included in the invention are cells transformed with a polynucleotide according to the invention and cells wherein a (hemi)cellulase according to the invention is genetically modified to enhance its activity and/or level of expression.



WO 2004/018662 A2

NOVEL (HEMI)CELLULASES AND USES THEREOF

Field of the invention

5 The invention relates to newly identified polynucleotide sequence comprising a gene that encodes a novel (hemi)cellulase from *Aspergillus niger*. The invention features the full length nucleotide sequence of the novel gene, the cDNA sequence comprising the full length coding sequence of the novel (hemi)cellulase as well as the amino acid sequence of the full-length (hemi)cellulase and functional equivalents thereof. The
10 invention also relates to methods of using these enzymes in industrial processes and methods of diagnosing fungal infections. Also included in the invention are cells transformed with a polynucleotide according to the invention and cells wherein a (hemi)cellulase according to the invention is genetically modified to enhance its activity and/or level of expression.

15

Background of the invention

Baked products such as bread are prepared from a dough which is usually made from the basic ingredients (wheat) flour, water and optionally salt. Depending on the baked
20 products, other ingredients added may be sugars, flavours etceteras. For leavened products, primarily baker's yeast is used next to chemical leavening systems such as a combination of an acid (generating compound) and bicarbonate.

In order to improve the handling properties of the dough and/or the final properties of the baked products there is a continuous effort to develop processing aids with improving
25 properties. Processing aids are defined herein as compounds that improve the handling properties of the dough and/or the final properties of the baked products. Dough properties that may be improved comprise machineability, gas retaining capability, reduced stickiness, elasticity, extensibility, moldability etcetera. Properties of the baked products that may be improved comprise loaf volume, crust crispiness, crumb texture and softness, flavour
30 relative staleness and shelf life. These dough and/or baked product improving processing aids can be divided into two groups: chemical additives and enzymes (also referred to as baking enzymes).

Yeast, enzymes and chemical additives are generally added separately to the dough. Yeast may be added as a liquid suspension, in a compressed form or as

- 2 -

active dry (ADY) or instant dry yeast (IDY). The difference between these yeast formulations is the water- and yeast dry matter content. Liquid yeast has a yeast dry matter content of less than 25% (w/v). Cream yeast is a particular form of liquid yeast and has a dry matter content between 17 and 23% (w/v). Compressed yeast has a yeast dry matter content between 25-35% (w/v) while the dry yeast formulations have a yeast dry matter content between 92-98% (w/v).

Enzymes may be added in a dry, e.g. granulated form or in liquid form. The chemical additives are in most cases added in powder form. Also, processing aid compositions which are tailored to specific baking applications, may be composed of a dedicated mixture of chemical additives and enzyme.

The preparation of a dough from the ingredients and processing aids described above is well known in the art and comprises mixing of said ingredients and processing aids and one or more moulding and fermentation steps.

The preparation of baked products from such doughs is also well known in the art and may comprise molding and shaping and further fermentation of the dough followed by baking at required temperatures and baking times.

Chemical additives with improving properties comprise oxidising agents such as ascorbic acid, bromate and azodicarbonate, reducing agents such as L-cysteine and glutathione, emulsifiers acting as dough conditioners such as diacetyl tartaric esters of mono/diglycerides (DATEM), sodium stearoyl lactylate (SSL) or calcium stearoyl lactylate (CSL), or acting as crumb softeners such as glycerol monostearate (GMS) etceteras, fatty materials such as triglycerides (fat) or lecithin and others.

As a result of a consumer-driven need to replace the chemical additives by more natural products, several baking enzymes have been developed with dough and/or baked product improving properties and which are used in all possible combinations depending on the specific baking application conditions. Suitable enzymes include starch degrading enzymes, arabinoxylan- and other hemicellulose degrading enzymes, cellulose degrading enzymes, oxidizing enzymes, fatty material splitting enzymes, protein degrading, modifying or crosslinking enzymes.

Starch degrading enzymes are for instance endo-acting enzymes such as alpha-amylase, maltogenic amylase, pullulanase or other debranching enzymes and exo-acting enzymes that cleave off glucose (amyloglucosidase), maltose (beta-amylase), maltotriose, maltotetraose and higher oligosaccharides.

- 3 -

Arabinoxylan- and other hemicellulose degrading enzymes are for instance xylanases, pentosanases, hemicellulase, arabinofuranosidase, glucanase and others.

Cellulose degrading enzymes are for instance cellulase, cellobiohydrolase and beta-glucosidase.

5 Oxidizing enzymes are for instance glucose oxidase, hexose oxidase, pyranose oxidase, sulfhydryl oxidase, lipoxygenase, laccase, polyphenol oxidases and others.

Fatty material splitting enzymes are for instance triacylglycerol lipases, phospholipases (such as A₁, A₂, B, C and D) and galactolipases.

10 Protein degrading, modifying or crosslinking enzymes are for instance endo-acting proteases (serine proteases, metalloproteases, aspartyl proteases, thiol proteases), exo-acting peptidases that cleave off one amino acid, or dipeptide, tripeptide etceteras from the N-terminal (aminopeptidases) or C-terminal (carboxypeptidases) ends of the polypeptide chain, asparagines or glutamine deamidating enzymes such as deamidase and peptidoglutaminase or crosslinking enzymes such as transglutaminase.

15 Baking enzymes may conveniently be produced in microorganisms. Microbial baking enzymes are available from a variety of sources; *Bacillus spec.* are a common source of bacterial enzymes, whereas fungal enzymes are commonly produced in *Aspergillus spec.*

20 Baking enzymes may be used in a manifold of baked goods. The term "baked goods" is herein defined as to comprise bread products such as tin bread, loaves of bread, French bread as well as rolls, cakes, pies, muffins, yeast raised and cake doughnuts and the like.

25 In the above processes, it is advantageous to use baking enzymes that are obtained by recombinant DNA techniques. Such recombinant enzymes have a number of advantages over their traditionally purified counterparts. Recombinant enzymes may be produced at a low cost price, high yield, free from contaminating agents like bacteria or viruses but also free from bacterial toxins or contaminating other enzyme activities.

Object of the invention

30

It is an object of the invention to provide novel polynucleotides encoding novel (hemi)cellulases with improved properties. A further object is to provide naturally and recombinantly produced (hemi)cellulases as well as recombinant strains producing these. Also fusion polypeptides are part of the invention as well as methods of making

- 4 -

and using the polynucleotides and polypeptides according to the invention.

It is also an object of the invention to provide novel (hemi)cellulases, which solve at least one of the above-mentioned problems or to provide novel (hemi)cellulases, which have one or more improved properties if used in dough and/or baked products, selected from the group of increased strength of the dough, increased elasticity of the dough, increased stability of the dough, reduced stickiness of the dough, improved extensibility of the dough, improved machineability of the dough, increased volume of the baked product, improved crumb structure of the baked product, improved softness of the baked product, improved flavour of the baked product, improved anti-staling of the baked product, improved colour of the baked product, improved crust of the baked product or which have a broad substrate specificity.

Summary of the invention

The invention provides for novel polynucleotides encoding novel (hemi)cellulases, in particular enzymes with alpha-1,2-mannosidase activity (EC 3.2.1.113), endo-1,4-beta-xylanase (EC 3.2.1.8), alpha-L-arabinofuranosidase (EC 3.2.1.55), exo-beta-1,3-glucanase (EC 3.2.1.58) and cellulase (EC 3.2.1.4).

More in particular, the invention provides for polynucleotides having a nucleotide sequence that hybridises preferably under highly stringent conditions to a sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20. Consequently, the invention provides nucleic acids that are more than 40% such as about 60%, preferably 65%, more preferably 70%, even more preferably 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% homologous to the sequences selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20.

In a more preferred embodiment the invention provides for such an isolated polynucleotide obtainable from a filamentous fungus, in particular *Aspergillus niger* is preferred.

In one embodiment, the invention provides for an isolated polynucleotide comprising a nucleic acid sequence encoding a polypeptide with an amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or functional equivalents thereof.

- 5 -

In a further preferred embodiment, the invention provides an isolated polynucleotide encoding at least one functional domain of a polypeptide selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or functional equivalents thereof.

5 In a preferred embodiment the invention provides a (hemi)cellulase gene selected from the group consisting of SEQ ID NO: 1, 4, 7, 10, 13, 16 and 19. In another aspect the invention provides a polynucleotide, preferably a cDNA encoding an *Aspergillus niger* (hemi)cellulase whose amino acid sequence is selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or variants or fragments of that
10 polypeptide. In a preferred embodiment the cDNA has a sequence selected from the group consisting of SEQ ID NO: 2, 5, 8, 11, 14, 17 and 20 or functional equivalents thereof.

In an even further preferred embodiment, the invention provides for a polynucleotide comprising the coding sequence of the polynucleotides according to the
15 invention, preferred is the polynucleotide sequence selected from the group consisting of SEQ ID NO: 2, 5, 8, 11, 14, 17 and 20.

The invention also relates to vectors comprising a polynucleotide sequence according to the invention and primers, probes and fragments that may be used to amplify or detect the DNA according to the invention.

20 In a further preferred embodiment, a vector is provided wherein the polynucleotide sequence according to the invention is functionally linked with regulatory sequences suitable for expression of the encoded amino acid sequence in a suitable host cell, such as *Aspergillus niger* or *Aspergillus oryzae*. The invention also provides methods for preparing polynucleotides and vectors according to the invention.

25 The invention also relates to recombinantly produced host cells that contain heterologous or homologous polynucleotides according to the invention.

In another embodiment, the invention provides recombinant host cells wherein the expression of a (hemi)cellulase according to the invention is significantly increased or wherein the activity of the (hemi)cellulase is increased.

30 In another embodiment the invention provides for a recombinantly produced host cell that contains heterologous or homologous polynucleotide according to the invention and wherein the cell is capable of producing a functional (hemi)cellulase according to the invention, preferably a cell capable of over-expressing the

- 6 -

(hemi)cellulase according to the invention, for example an *Aspergillus* strain comprising an increased copy number of a gene or cDNA according to the invention.

In yet another aspect of the invention, a purified polypeptide is provided. The polypeptides according to the invention include the polypeptides encoded by the polynucleotides according to the invention. Especially preferred is a polypeptide selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or functional equivalents thereof.

Accordingly, in one aspect the present invention provides a (hemi)cellulase composition containing as an active ingredient an enzyme selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or functional equivalents thereof.

In another aspect, the invention provides a method of making baked goods wherein there is incorporated into the dough used for making the baked goods one or more enzymes selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or functional equivalents thereof.

Fusion proteins comprising a polypeptide according to the invention are also within the scope of the invention. The invention also provides methods of making the polypeptides according to the invention.

The invention also relates to the use of the (hemi)cellulase according to the invention in any industrial process as described herein.

Detailed description of the invention

Polynucleotides

The present invention provides polynucleotides encoding (hemi)cellulases having an amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or functional equivalents thereof. The sequences of the seven genes encoding the (hemi)cellulases NBE012, NBE021, NBE022, NBE064, NBE075, NBE092 and NBE097 respectively were determined by sequencing genomic clones obtained from *Aspergillus niger*. The invention provides polynucleotide sequences comprising the genes encoding the (hemi)cellulase NBE012, NBE021, NBE022, NBE064, NBE075, NBE092 and NBE097 as well as their complete cDNA sequences and their coding sequences (Table 1). Accordingly, the invention relates to isolated polynucleotides comprising the nucleotide sequences selected from the group consisting

- 7 -

of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20 or functional equivalents thereof.

Table 1.

hemi(cellulase)			Sequence (SEQ ID NO)		
NBE xxx	EC number	enzyme activity	genomic	cDNA	amino acid
NBE012	EC 3.2.1.113	alpha-1,2-mannosidase	1	2	3
NBE021	EC 3.2.1.8	endo-1,4-beta-xylanase	4	5	6
NBE022	EC 3.2.1.8	endo-1,4-beta-xylanase	7	8	9
NBE064	EC 3.2.1.55	alpha-L-arabinofuranosidase	10	11	12
NBE075	EC 3.2.1.4	cellulase	13	14	15
NBE092	EC 3.2.1.4	cellulase	16	17	18
NBE097	EC 3.2.1.58	exo-beta-1,3-glucanase	19	20	21

5

More in particular, the invention relates to an isolated polynucleotide hybridisable under stringent conditions, preferably under highly stringent conditions, to a polynucleotide selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20. Advantageously, such polynucleotides may be obtained from filamentous fungi, in particular from *Aspergillus niger*. More specifically, the invention relates to an isolated polynucleotide having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20.

10

The invention also relates to an isolated polynucleotide encoding at least one functional domain of a polypeptide having an amino acid sequences selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or functional equivalents thereof.

15

As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which may be isolated from chromosomal DNA, which include an open reading frame encoding a protein, e.g. an *Aspergillus niger* (hemi)cellulase. A gene may include coding sequences, non-coding sequences, introns and regulatory sequences. Moreover, a gene refers to an isolated nucleic acid molecule as defined herein.

20

A nucleic acid molecule of the present invention, such as a nucleic acid molecule having the nucleotide sequence selected from the group consisting of SEQ ID

NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20 or a functional equivalent thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, using all or portion of the nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13,
5 14, 16, 17, 19 and 20 as a hybridization probe, nucleic acid molecules according to the invention can be isolated using standard hybridization and cloning techniques (e. g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

10 Moreover, a nucleic acid molecule encompassing all or a portion of the nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20 can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence information contained in the nucleic acid sequence selected from the group consisting of SEQ ID
15 NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis.

20 Furthermore, oligonucleotides corresponding to or hybridisable to nucleotide sequences according to the invention can be prepared by standard synthetic techniques, e. g., using an automated DNA synthesizer.

In one preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO: 2. The sequence of SEQ ID
25 NO: 2 corresponds to the coding region of the *Aspergillus niger* gene provided in SEQ ID NO: 1. This cDNA comprises the sequence encoding the *Aspergillus niger* NBE012 polypeptide as shown in SEQ ID NO: 3.

In a second preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO: 5. The sequence of
30 SEQ ID NO: 5 corresponds to the coding region of the *Aspergillus niger* gene provided in SEQ ID NO: 4. This cDNA comprises the sequence encoding the *Aspergillus niger* NBE021 polypeptide as shown in SEQ ID NO: 6.

In a third preferred embodiment, an isolated nucleic acid molecule of the

- 9 -

invention comprises the nucleotide sequence shown in SEQ ID NO: 8. The sequence of SEQ ID NO: 8 corresponds to the coding region of the *Aspergillus niger* gene provided in SEQ ID NO: 7. This cDNA comprises the sequence encoding the *Aspergillus niger* NBE022 polypeptide as shown in SEQ ID NO: 9.

5 In a fourth preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO: 11. The sequence of SEQ ID NO: 11 corresponds to the coding region of the *Aspergillus niger* gene provided in SEQ ID NO: 10. This cDNA comprises the sequence encoding the *Aspergillus niger* NBE064 polypeptide as shown in SEQ ID NO: 12.

10 In a fifth preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO: 14. The sequence of SEQ ID NO: 14 corresponds to the coding region of the *Aspergillus niger* gene provided in SEQ ID NO: 13. This cDNA comprises the sequence encoding the *Aspergillus niger* NBE075 polypeptide as shown in SEQ ID NO: 15.

15 In a sixth preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO: 17. The sequence of SEQ ID NO: 17 corresponds to the coding region of the *Aspergillus niger* gene provided in SEQ ID NO: 16. This cDNA comprises the sequence encoding the *Aspergillus niger* NBE092 polypeptide as shown in SEQ ID NO: 18.

20 In a seventh preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO: 20. The sequence of SEQ ID NO: 20 corresponds to the coding region of the *Aspergillus niger* gene provided in SEQ ID NO: 19. This cDNA comprises the sequence encoding the *Aspergillus niger* NBE097 polypeptide as shown in SEQ ID NO: 21.

25 In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20 or a functional equivalent of these nucleotide sequences. A nucleic acid molecule, which is complementary to another nucleotide sequence, is one that is
30 sufficiently complementary to the other nucleotide sequence such that it can hybridize to the other nucleotide sequence thereby forming a stable duplex.

One aspect of the invention pertains to isolated nucleic acid molecules that encode a polypeptide of the invention or a functional equivalent thereof such as a

- 10 -

biologically active fragment or domain, as well as nucleic acid molecules sufficient for use as hybridisation probes to identify nucleic acid molecules encoding a polypeptide of the invention and fragments of such nucleic acid molecules suitable for use as PCR primers for the amplification or mutation of nucleic acid molecules. An "isolated polynucleotide" or "isolated nucleic acid" is a DNA or RNA that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (one on the 5' end and one on the 3' end) in the naturally occurring genome of the organism from which it is derived. Thus, in one embodiment, an isolated nucleic acid includes some or all of the 5' non-coding (e.g., promotor) sequences that are immediately contiguous to the coding sequence. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA that is part of a hybrid gene encoding an additional polypeptide that is substantially free of cellular material, viral material, or culture medium (when produced by recombinant DNA techniques), or chemical precursors or other chemicals (when chemically synthesized). Moreover, an "isolated nucleic acid fragment" is a nucleic acid fragment that is not naturally occurring as a fragment and would not be found in the natural state.

As used herein, the terms "polynucleotide" or "nucleic acid molecule" are intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. The nucleic acid may be synthesized using oligonucleotide analogs or derivatives (e.g., inosine or phosphorothioate nucleotides). Such oligonucleotides can be used, for example, to prepare nucleic acids that have altered base-pairing abilities or increased resistance to nucleases.

Another embodiment of the invention provides an isolated nucleic acid molecule which is antisense to a nucleic acid molecule according to the invention. Also included within the scope of the invention are the complement strands of the nucleic acid molecules described herein.

- 11 -

Sequencing errors

The sequence information as provided herein should not be so narrowly construed as to require inclusion of erroneously identified bases. The specific sequences disclosed herein can be readily used to isolate the complete gene from filamentous fungi, in particular *Aspergillus niger* which in turn can easily be subjected to further sequence analyses thereby identifying sequencing errors.

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

The person skilled in the art is capable of identifying such erroneously identified bases and knows how to correct for such errors.

Nucleic acid fragments, probes and primers

A nucleic acid molecule according to the invention may comprise only a portion or a fragment of the nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20, for example a fragment which can be used as a probe or primer or a fragment encoding a portion of protein according to the invention. The nucleotide sequence determined from the cloning of the (hemi)cellulase gene and cDNA allows for the generation of probes and primers

- 12 -

designed for use in identifying and/or cloning other (hemi)cellulase family members, as well as (hemi)cellulase homologues from other species. The probe/primer typically comprises substantially purified oligonucleotide which typically comprises a region of nucleotide sequence that hybridizes preferably under highly stringent conditions to at least about 12 or 15, preferably about 18 or 20, preferably about 22 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 or more consecutive nucleotides of a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20 or of a functional equivalent thereof.

Probes based on the nucleotide sequences provided herein can be used to detect transcripts or genomic sequences encoding the same or homologous proteins for instance in other organisms. In preferred embodiments, the probe further comprises a label group attached thereto, e.g., the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme cofactor. Such probes can also be used as part of a diagnostic test kit for identifying cells that express a (hemi)cellulase protein.

Identity & homology

The terms "homology" or "percent identity" are used interchangeably herein. For the purpose of this invention, it is defined here that in order to determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = number of identical positions/total number of positions (i.e. overlapping positions) x 100). Preferably, the two sequences are the same length.

The skilled person will be aware of the fact that several different computer programmes are available to determine the homology between two sequences. For instance, a comparison of sequences and determination of percent identity between two

- 13 -

sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (J. Mol. Biol. (48): 444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. The skilled person will appreciate that all these different parameters will yield slightly different results but that the overall percentage identity of two sequences is not significantly altered when using different algorithms.

10 In yet another embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity two amino acid or nucleotide sequence is determined using the
15 algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989) which has been incorporated into the ALIGN program (version 2.0) (available at: <http://vega.igh.cnrs.fr/bin/align-guess.cgi>) using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be
20 used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the BLASTN and BLASTX programs (version 2.0) of Altschul, et al. (1990) J. Mol. Biol. 215:403—10. BLAST nucleotide searches can be performed with the BLASTN program, score = 100, word length = 12 to obtain nucleotide sequences
25 homologous to PLP03 nucleic acid molecules of the invention. BLAST protein searches can be performed with the BLASTX program, score = 50, word length = 3 to obtain amino acid sequences homologous to PLP03 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) Nucleic Acids Res. 25(17):3389-3402. When utilizing
30 BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., BLASTX and BLASTN) can be used. See <http://www.ncbi.nlm.nih.gov>.

- 14 -

Hybridisation

As used herein, the term "hybridizing" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least about 50%, at least about 60%, at least about 70%, more preferably at least about 80%, even more preferably at least about 85% to 90%, more preferably at least 95% homologous to each other typically remain hybridized to each other.

A preferred, non-limiting example of such hybridization conditions are hybridization in 6x sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 1x SSC, 0.1 % SDS at 50°C, preferably at 55°C, preferably at 60°C and even more preferably at 65°C.

Highly stringent conditions include, for example, hybridizing at 68°C in 5x SSC/5x Denhardt's solution / 1.0% SDS and washing in 0.2x SSC/0.1% SDS at room temperature. Alternatively, washing may be performed at 42°C.

The skilled artisan will know which conditions to apply for stringent and highly stringent hybridisation conditions. Additional guidance regarding such conditions is readily available in the art, for example, in Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; and Ausubel et al. (eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, N.Y.).

Of course, a polynucleotide which hybridizes only to a poly A sequence (such as the 3' terminal poly(A) tract of mRNAs), or to a complementary stretch of T (or U) residues, would not be included in a polynucleotide of the invention used to specifically hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

Obtaining full length DNA from other organisms

In a typical approach, cDNA libraries constructed from other organisms, e.g. filamentous fungi, in particular from the species *Aspergillus* can be screened.

For example, *Aspergillus* strains can be screened for homologous polynucleotides by Northern blot analysis. Upon detection of transcripts homologous to polynucleotides according to the invention, cDNA libraries can be constructed from RNA isolated from the appropriate strain, utilizing standard techniques well known to those of

- 15 -

skill in the art. Alternatively, a total genomic DNA library can be screened using a probe hybridisable to a polynucleotide according to the invention.

Homologous gene sequences can be isolated, for example, by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of nucleotide sequences as taught herein.

The template for the reaction can be cDNA obtained by reverse transcription of mRNA prepared from strains known or suspected to express a polynucleotide according to the invention. The PCR product can be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a new PLP03 nucleic acid sequence, or a functional equivalent thereof.

The PCR fragment can then be used to isolate a full length cDNA clone by a variety of known methods. For example, the amplified fragment can be labeled and used to screen a bacteriophage or cosmid cDNA library. Alternatively, the labeled fragment can be used to screen a genomic library.

PCR technology can also be used to isolate full-length cDNA sequences from other organisms. For example, RNA can be isolated, following standard procedures, from an appropriate cellular or tissue source. A reverse transcription reaction can be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis.

The resulting RNA/DNA hybrid can then be "tailed" (e.g., with guanines) using a standard terminal transferase reaction, the hybrid can be digested with RNase H, and second strand synthesis can then be primed (e.g., with a poly-C primer). Thus, cDNA sequences upstream of the amplified fragment can easily be isolated. For a review of useful cloning strategies, see e.g., Sambrook et al., supra; and Ausubel et al., supra.

Vectors

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a protein according to the invention or a functional equivalent thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector,

- 16 -

wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated
5 along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. The terms "plasmid" and "vector" can
10 be used interchangeably herein as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of
15 the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vector includes one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operatively linked" is intended to mean that the nucleotide sequence
20 of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signal). Such regulatory sequences are
25 described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive or inducible expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in a certain host cell (e.g. tissue-specific regulatory sequences). It will be
30 appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, encoded by nucleic

- 17 -

acids as described herein (e.g. (hemi)cellulases, mutant (hemi)cellulases, fragments thereof, variants or functional equivalents thereof, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of (hemi)cellulases in prokaryotic or eukaryotic cells. For example, a protein according to the invention can be expressed in bacterial cells such as *E. coli* and *Bacillus species*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression vectors useful in the present invention include chromosomal-, episomal- and virus-derived vectors e.g., vectors derived from bacterial plasmids, bacteriophage, yeast episome, yeast chromosomal elements, viruses such as baculoviruses, papova viruses, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. The skilled person will know other suitable promoters. In a specific embodiment, promoters are preferred that are capable of directing a high expression level of (hemi)cellulases in filamentous fungi. Such promoters are known in the art. The expression constructs may contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, transduction, infection, lipofection, cationic lipid-mediated transfection or

- 18 -

electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*Molecular Cloning: A Laboratory Manual*, 2nd, ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), Davis et al., *Basic Methods in Molecular Biology* (1986) and other laboratory manuals.

5 For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred
10 selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methatrexate. A nucleic acid encoding a selectable marker is preferably introduced into a host cell on the same vector as that encoding a protein according to the invention or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g. cells that have
15 incorporated the selectable marker gene will survive, while the other cells die).

Expression of proteins in prokaryotes is often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, e.g. to the amino terminus of the recombinant protein. Such fusion vectors
20 typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from
25 the fusion moiety after purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase.

As indicated, the expression vectors will preferably contain selectable markers. Such markers include dihydrofolate reductase or neomycin resistance for eukarotic cell culture and tetracycline or ampicilling resistance for culturing in *E. coli* and other bacteria.
30 Representative examples of appropriate host include bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium*; fungal cells, such as yeast; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9; animal cells such as CHO, COS and Bowes melanoma; and plant cells. Appropriate culture mediums and conditions for the above-

- 19 -

described host cells are known in the art.

Among vectors preferred for use in bacteria are pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16A, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic
5 vectors are PWLNEO, pSV2CAT, pOG44, pZT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Known bacterial promoters for use in the present invention include *E. coli* lacI
10 and lacZ promoters, the T3 and T7 promoters, the gpt promoter, the lambda PR, PL promoters and the trp promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter.

15 Inserting an enhancer sequence into the vector may increase transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication
20 origin at bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signal may be incorporated into the expressed polypeptide. The signals may
25 be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and
30 persistence in the host cell, during purification or during subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification.

Polypeptides according to the invention

The invention provides an isolated polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21, an amino acid sequence obtainable by expressing the polynucleotide selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20 in an appropriate host. Also, a peptide or polypeptide comprising a functional equivalent of the above polypeptides is comprised within the present invention. The above polypeptides are collectively comprised in the term "polypeptides according to the invention"

The terms "peptide" and "oligopeptide" are considered synonymous (as is commonly recognized) and each term can be used interchangeably as the context requires to indicate a chain of at least two amino acids coupled by peptidyl linkages. The word "polypeptide" is used herein for chains containing more than seven amino acid residues. All oligopeptide and polypeptide formulas or sequences herein are written from left to right and in the direction from amino terminus to carboxy terminus. The one-letter code of amino acids used herein is commonly known in the art and can be found in Sambrook, et al. (*Molecular Cloning: A Laboratory Manual*, 2nd, ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989)

By "isolated" polypeptide or protein is intended a polypeptide or protein removed from its native environment. For example, recombinantly produced polypeptides and proteins expressed in host cells are considered isolated for the purpose of the invention as are native or recombinant polypeptides which have been substantially purified by any suitable technique such as, for example, the single-step purification method disclosed in Smith and Johnson, *Gene* 67:31-40 (1988).

The (hemi)cellulase according to the invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant

- 21 -

techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

A (hemi)cellulase according to the invention may be advantageously used in baking processes. The amount of enzyme to be added to the dough is determined empirically. It may depend on the quality of the flour used, the degree of improvement which is required, the kind of bread or baked goods, the method of preparing the dough, the proportion of other ingredients etcetera.

Protein fragments

The invention also features biologically active fragments of the polypeptides according to the invention.

Biologically active fragments of a polypeptide of the invention include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the (hemi)cellulase (e.g., the amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21), which include fewer amino acids than the full length protein, and exhibit at least one biological activity of the corresponding full-length protein. Typically, biologically active fragments comprise a domain or motif with at least one activity of the corresponding full length protein.

A biologically active fragment of a protein of the invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the biological activities of the native form of a polypeptide of the invention.

The invention also features nucleic acid fragments that encode the above biologically active fragments of the (hemi)cellulase protein.

Fusion proteins

The proteins of the present invention or functional equivalents thereof, e.g.,

- 22 -

biologically active portions thereof, can be operatively linked to a non-(hemi)cellulase polypeptide (e.g., heterologous amino acid sequences) to form fusion proteins. As used herein, a (hemi)cellulase "chimeric protein" or "fusion protein" comprises a (hemi)cellulase polypeptide operatively linked to a non-(hemi)cellulase polypeptide. A
5 "(hemi)cellulase polypeptide" refers to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21, whereas a "non-(hemi)cellulase polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the (hemi)cellulase, e.g., a protein which is different from the (hemi)cellulase and which is
10 derived from the same or a different organism. Within a (hemi)cellulase fusion protein the (hemi)cellulase polypeptide can correspond to all or a portion of a (hemi)cellulase protein. In a preferred embodiment, a (hemi)cellulase fusion protein comprises at least one biologically active fragment of a (hemi)cellulase protein. In another preferred embodiment, a (hemi)cellulase fusion protein comprises at least two biologically active
15 portions of a (hemi)cellulase protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the (hemi)cellulase polypeptide and the non-(hemi)cellulase polypeptide are fused in-frame to each other. The non-(hemi)cellulase polypeptide can be fused to the N-terminus or C-terminus of the (hemi)cellulase polypeptide.

20 For example, in one embodiment, the fusion protein is a GST-(hemi)cellulase fusion protein in which the (hemi)cellulase sequence is fused to the C-terminus of the GST sequence. Such fusion proteins can facilitate the purification of recombinant (hemi)cellulase(s). In another embodiment, the fusion protein is a (hemi)cellulase protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g.,
25 mammalian and Yeast host cells), expression and/or secretion of (hemi)cellulase can be increased through use of a heterologous signal sequence.

In another example, the gp67 secretory sequence of the baculovirus envelope protein can be used as a heterologous signal sequence (*Current Protocols in Molecular Biology*, Ausubel et al., eds., John Wiley & Sons, 1992). Other examples of eukaryotic
30 heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase (Stratagene; La Jolla, California). In yet another example, useful prokaryotic heterologous signal sequences include the phoA secretory signal (Sambrook et al., *supra*) and the protein A secretory signal (Pharmacia Biotech;

- 23 -

Piscataway, New Jersey).

A signal sequence can be used to facilitate secretion and isolation of a protein or polypeptide of the invention. Signal sequences are typically characterized by a core of hydrophobic amino acids which are generally cleaved from the mature protein during secretion in one or more cleavage events. Such signal peptides contain processing sites that allow cleavage of the signal sequence from the mature proteins as they pass through the secretory pathway. The signal sequence directs secretion of the protein, such as from a eukaryotic host into which the expression vector is transformed, and the signal sequence is subsequently or concurrently cleaved. The protein can then be readily purified from the extracellular medium by art recognized methods. Alternatively, the signal sequence can be linked to the protein of interest using a sequence which facilitates purification, such as with a GST domain. Thus, for instance, the sequence encoding the polypeptide may be fused to a marker sequence, such as a sequence encoding a peptide, which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (Qiagen, Inc.), among others, many of which are commercially available. As described in *Gentz et al, Proc. Natl. Acad. Sci. USA* 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. The HA tag is another peptide useful for purification which corresponds to an epitope derived of influenza hemagglutinin protein, which has been described by Wilson *et al.*, *Cell* 37:767 (1984), for instance.

Preferably, a chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley &

- 24 -

Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g. a GST polypeptide). A (hemi)cellulase-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the (hemi)cellulase protein.

5

Functional equivalents

The terms "functional equivalents" and "functional variants" are used interchangeably herein. Functional equivalents of (hemi)cellulase encoding DNA are isolated DNA fragments that encode a polypeptide that exhibits a particular function of the *Aspergillus niger* (hemi)cellulase as defined herein. A functional equivalent of a (hemi)cellulase polypeptide according to the invention is a polypeptide that exhibits at least one function of an *Aspergillus niger* (hemi)cellulase as defined herein. Functional equivalents therefore also encompass biologically active fragments.

10
15
20
25
Functional protein or polypeptide equivalents may contain only conservative substitutions of one or more amino acids in the amino acid sequences selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or substitutions, insertions or deletions of non-essential amino acids. Accordingly, a non-essential amino acid is a residue that can be altered in the amino acid sequences selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 without substantially altering the biological function. For example, amino acid residues that are conserved among the (hemi)cellulase proteins of the present invention, are predicted to be particularly unamenable to alteration. Furthermore, amino acids conserved among the (hemi)cellulase proteins according to the present invention and other (hemi)cellulases are not likely to be amenable to alteration.

30
The term "conservative substitution" is intended to mean that a substitution in which the amino acid residue is replaced with an amino acid residue having a similar side chain. These families are known in the art and include amino acids with basic side chains (e.g. lysine, arginine and histidine), acidic side chains (e.g. aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), non-polar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine,

- 25 -

phenylalanine tryptophan, histidine).

Functional nucleic acid equivalents may typically contain silent mutations or mutations that do not alter the biological function of encoded polypeptide. Accordingly, the invention provides nucleic acid molecules encoding (hemi)cellulase proteins that
5 contain changes in amino acid residues that are not essential for a particular biological activity. Such (hemi)cellulase proteins differ in amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 yet retain at least one biological activity. In one embodiment the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises a substantially
10 homologous amino acid sequence of at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous to the amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21.

For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J.U. et al., Science 247:1306-1310 (1990)
15 wherein the authors indicate that there are two main approaches for studying the tolerance of an amino acid sequence to change. The first method relies on the process of evolution, in which mutations are either accepted or rejected by natural selection. The second approach uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene and selects or screens to identify sequences that maintain
20 functionality. As the authors state, these studies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which changes are likely to be permissive at a certain position of the protein. For example, most buried amino acid residues require non-polar side chains, whereas few features of surface side chains are generally conserved. Other such phenotypically silent
25 substitutions are described in Bowie et al, supra, and the references cited therein.

An isolated nucleic acid molecule encoding a protein homologous to the protein selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 can be created by introducing one or more nucleotide substitutions, additions or deletions into the corresponding coding nucleotide sequences (Table 1) such that one or more amino
30 acid substitutions, deletions or insertions are introduced into the encoded protein. Such mutations may be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis.

The term "functional equivalents" also encompasses orthologues of the

- 26 -

Aspergillus niger (hemi)cellulases provided herein. Orthologues of the *Aspergillus niger* (hemi)cellulases are proteins that can be isolated from other strains or species and possess a similar or identical biological activity. Such orthologues can readily be identified as comprising an amino acid sequence that is substantially homologous to the amino acid sequences selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21.

As defined herein, the term "substantially homologous" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., with similar side chain) amino acids or nucleotides to a second amino acid or nucleotide sequence such that the first and the second amino acid or nucleotide sequences have a common domain. For example, amino acid or nucleotide sequences which contain a common domain having about 60%, preferably 65%, more preferably 70%, even more preferably 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identity or more are defined herein as sufficiently identical.

Also, nucleic acids encoding other (hemi)cellulase family members, which thus have a nucleotide sequence that differs from a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20, are within the scope of the invention. Moreover, nucleic acids encoding (hemi)cellulase proteins from different species which thus have a nucleotide sequence which differs from a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20 are within the scope of the invention.

Nucleic acid molecules corresponding to variants (e.g. natural allelic variants) and homologues of the polynucleotides of the invention can be isolated based on their homology to the nucleic acids disclosed herein using the cDNAs disclosed herein or a suitable fragment thereof, as a hybridisation probe according to standard hybridisation techniques preferably under highly stringent hybridisation conditions.

In addition to naturally occurring allelic variants of the *Aspergillus niger* sequences provided herein, the skilled person will recognise that changes can be introduced by mutation into the nucleotide sequences selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20 thereby leading to changes in the amino acid sequence of the (hemi)cellulase protein without substantially altering the function of the protein.

In another aspect of the invention, improved (hemi)cellulase are provided.

- 27 -

Improved (hemi)cellulases are proteins wherein at least one biological activity is improved. Such proteins may be obtained by randomly introducing mutations along all or part of the (hemi)cellulase coding sequence, such as by saturation mutagenesis, and the resulting mutants can be expressed recombinantly and screened for biological activity.

5 For instance, the art provides for standard assays for measuring the enzymatic activity of (hemi)cellulases and thus improved proteins may easily be selected.

In a preferred embodiment the (hemi)cellulase has an amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21. In another embodiment, the (hemi)cellulase is substantially homologous to the amino acid
10 sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 and retains at least one biological activity of a polypeptide selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21, yet differs in amino acid sequence due to natural variation or mutagenesis as described above.

In a further preferred embodiment, the (hemi)cellulase has an amino acid
15 sequence encoded by an isolated nucleic acid fragment capable of hybridising to a nucleic acid selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20, preferably under highly stringent hybridisation conditions.

Accordingly, the (hemi)cellulase is a protein which comprises an amino acid sequence at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%,
20 99% or more homologous to the amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 and retains at least one functional activity of the polypeptide selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21.

In particular, the (hemi)cellulase is a protein which comprises an amino acid
25 sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous to the amino acid sequence shown in SEQ ID NO: 3 or the (hemi)cellulase is a protein which comprises an amino acid sequence at least about 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous to the amino acid sequence shown in SEQ ID NO: 6, or the (hemi)cellulase
30 is a protein which comprises an amino acid sequence at least about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous to the amino acid sequence shown in SEQ ID NO: 9, or the (hemi)cellulase is a protein which comprises an amino acid sequence at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%,

- 28 -

90%, 95%, 96%, 97%, 98%, 99% or more homologous to the amino acid sequence shown in SEQ ID NO: 12 or the (hemi)cellulase is a protein which comprises an amino acid sequence at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous to the amino acid sequence shown in SEQ ID NO: 15, or
5 the (hemi)cellulase is a protein which comprises an amino acid sequence at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous to the amino acid sequence shown in SEQ ID NO: 18 or the (hemi)cellulase is a protein which comprises an amino acid sequence at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more
10 homologous to the amino acid sequence shown in SEQ ID NO: 21.

Functional equivalents of a protein according to the invention can also be identified e.g. by screening combinatorial libraries of mutants, e.g. truncation mutants, of the protein of the invention for (hemi)cellulase activity. In one embodiment, a variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level. A
15 variegated library of variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential protein sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display). There are a variety of methods that can be used to produce libraries of potential variants of the polypeptides of the
20 invention from a degenerate oligonucleotide sequence. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al. (1984) Science 198:1056; Ike et al. (1983) Nucleic Acid Res. 11:477).

In addition, libraries of fragments of the coding sequence of a polypeptide of the
25 invention can be used to generate a variegated population of polypeptides for screening a subsequent selection of variants. For example, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of the coding sequence of interest with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double
30 stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal

- 29 -

fragments of various sizes of the protein of interest.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations of truncation, and for screening cDNA libraries for gene products having a selected property. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify variants of a protein of the invention (Arkin and Yourvan (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331).

It will be apparent for the person skilled in the art that DNA sequence polymorphisms that may lead to changes in the amino acid sequence of the (hemi)cellulase may exist within a given population. Such genetic polymorphisms may exist in cells from different populations or within a population due to natural allelic variation. Allelic variants may also include functional equivalents.

Fragments of a polynucleotide according to the invention may also comprise polynucleotides not encoding functional polypeptides. Such polynucleotides may function as probes or primers for a PCR reaction.

Nucleic acids according to the invention irrespective of whether they encode functional or non-functional polypeptides, can be used as hybridization probes or polymerase chain reaction (PCR) primers. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having a (hemi)cellulase activity include, inter alia, (1) isolating the gene encoding the (hemi)cellulase protein, or allelic variants thereof from a cDNA library e.g. from other organisms than *Aspergillus niger*; (2) in situ hybridization (e.g. FISH) to metaphase chromosomal spreads to provide precise chromosomal location of the (hemi)cellulase gene as described in Verma et al., Human Chromosomes: a Manual of Basic Techniques, Pergamon Press, New York (1988); (3) Northern blot analysis for detecting expression of (hemi)cellulase mRNA in specific tissues and/or cells and 4) probes and primers that can be used as a diagnostic tool to analyse the presence of a nucleic acid hybridisable to the (hemi)cellulase probe in a

- 30 -

given biological (e.g. tissue) sample.

Also encompassed by the invention is a method of obtaining a functional equivalent of a (hemi)cellulase encoding gene or cDNA. Such a method entails obtaining a labelled probe that includes an isolated nucleic acid which encodes all or a portion of the sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or a variant thereof; screening a nucleic acid fragment library with the labelled probe under conditions that allow hybridisation of the probe to nucleic acid fragments in the library, thereby forming nucleic acid duplexes, and preparing a full-length gene sequence from the nucleic acid fragments in any labelled duplex to obtain a gene related to the (hemi)cellulase gene.

In one embodiment, a nucleic acid of the invention is at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more homologous to a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20 or the complement thereof.

In another preferred embodiment a polypeptide of the invention is at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more homologous to the amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21.

Host cells

In another embodiment, the invention features cells, e.g., transformed host cells or recombinant host cells that contain a nucleic acid encompassed by the invention. A "transformed cell" or "recombinant cell" is a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid according to the invention. Both prokaryotic and eukaryotic cells are included, e.g., bacteria, fungi, yeast, and the like, especially preferred are cells from filamentous fungi, in particular *Aspergillus niger*.

A host cell can be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in a specific, desired fashion. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may facilitate optimal functioning of the protein.

- 31 -

Various host cells have characteristic and specific mechanisms for post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems familiar to those of skill in the art of molecular biology and/or microbiology can be chosen to ensure the desired and correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product can be used. Such host cells are well known in the art.

Host cells also include, but are not limited to, mammalian cell lines such as CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, and choroid plexus cell lines.

If desired, the polypeptides according to the invention can be produced by a stably-transfected cell line. A number of vectors suitable for stable transfection of mammalian cells are available to the public, methods for constructing such cell lines are also publicly known, e.g., in Ausubel et al. (supra).

Antibodies

The invention further features antibodies, such as monoclonal or polyclonal antibodies, that specifically bind (hemi)cellulase proteins according to the invention.

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to (hemi)cellulase protein. Fab and F(ab')₂ fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding of an intact antibody (Wahl *et al.*, *J. Nucl. Med.* 24:316-325 (1983)). Thus, these fragments are preferred.

The antibodies of the present invention may be prepared by any of a variety of methods. For example, cells expressing the (hemi)cellulase protein or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of (hemi)cellulase protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

- 32 -

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or (hemi)cellulase protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology (Kohler *et al.*, *Nature* 256:495 (1975); Kohler *et al.*, *Eur. J. Immunol.* 6:511 (1976); Hammerling *et al.*,
5 In: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981)). In general, such procedures involve immunizing an animal (preferably a mouse) with a (hemi)cellulase protein antigen or, with a (hemi)cellulase protein expressing cell. The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention;
10 however, it is preferably to employ the parent myeloma cell line (SP₂O), available from the American Type Culture Collection, Rockville, Maryland. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands *et al.* (*Gastro-enterology* 80:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones
15 which secrete antibodies capable of binding the (hemi)cellulase protein antigen. In general, the polypeptides can be coupled to a carrier protein, such as KLH, as described in Ausubel *et al.*, *supra*, mixed with an adjuvant, and injected into a host mammal.

In particular, various host animals can be immunized by injection of a polypeptide of interest. Examples of suitable host animals include rabbits, mice, guinea
20 pigs, and rats. Various adjuvants can be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), adjuvant mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, BCG (bacille Calmette-Guerin) and
25 *Corynebacterium parvum*. Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of the immunized animals.

Such antibodies can be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD, and any subclass thereof. The hybridomas producing the mAbs of this invention can be cultivated *in vitro* or *in vivo*.

30 Once produced, polyclonal or monoclonal antibodies are tested for specific recognition of a protein according to the invention or functional equivalent thereof in an immunoassay, such as a Western blot or immunoprecipitation analysis using standard techniques, e.g., as described in Ausubel *et al.*, *supra*. Antibodies that specifically bind

- 33 -

to a protein according to the invention or functional equivalents thereof are useful in the invention. For example, such antibodies can be used in an immunoassay to detect a protein according to the invention in pathogenic or non-pathogenic strains of *Aspergillus* (e.g., in *Aspergillus* extracts).

5 Preferably, antibodies of the invention are produced using fragments of a protein according to the invention that appear likely to be antigenic, by criteria such as high frequency of charged residues. For example, such fragments may be generated by standard techniques of PCR, and then cloned into the pGEX expression vector (Ausubel et al., supra). Fusion proteins may then be expressed in *E. coli* and purified using a
10 glutathione agarose affinity matrix as described in Ausubel, et al., supra. If desired, several (e.g., two or three) fusions can be generated for each protein, and each fusion can be injected into at least two rabbits. Antisera can be raised by injections in a series, typically including at least three booster injections. Typically, the antisera are checked for their ability to immunoprecipitate the protein according to the invention or functional
15 equivalents thereof whereas unrelated proteins may serve as a control for the specificity of the immune reaction.

 Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778 and 4,704,692) can be adapted to produce single chain antibodies against a protein according to the invention or functional equivalents thereof.
20 Kits for generating and screening phage display libraries are commercially available e.g. from Pharmacia.

 Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223, 409; PCT Publication No. WO 92/18619; PCT Publication No. WO
25 91/17271; PCT Publication No. WO 20791; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs et al. (1991) *Bio/Technology* 9:1370-1372; Hay et al. (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse et al. (1989) *Science* 246:1275-1281; Griffiths et al. (1993)
30 *EMBO J.* 12:725-734.

 Polyclonal and monoclonal antibodies that specifically bind a protein according to the invention or functional equivalents thereof can be used, for example, to detect expression of a gene encoding a protein according to the invention or a functional

- 34 -

equivalent thereof e.g. in another strain of *Aspergillus*. For example, a protein according to the invention can be readily detected in conventional immunoassays of *Aspergillus* cells or extracts. Examples of suitable assays include, without limitation, Western blotting, ELISAs, radioimmune assays, and the like.

5 By "specifically binds" is meant that an antibody recognizes and binds a particular antigen, e.g., a protein according to the invention, but does not substantially recognize and bind other unrelated molecules in a sample.

Antibodies can be purified, for example, by affinity chromatography methods in which the polypeptide antigen is immobilized on a resin.

10 An antibody directed against a polypeptide of the invention (e.g., monoclonal antibody) can be used to isolate the polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, such an antibody can be used to detect the protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the polypeptide. The antibodies can also be
15 used diagnostically to monitor protein levels in cells or tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen or in the diagnosis of Aspergillosis..

Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups,
20 fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a
25 luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive materials include ^{125}I , ^{131}I , ^{35}S or ^3H .

Preferred epitopes encompassed by the antigenic peptide are regions that are located on the surface of the protein, e.g., hydrophilic regions. Hydrophobicity plots of
30 the proteins of the invention can be used to identify hydrophilic regions.

The antigenic peptide of a protein of the invention comprises at least 7 (preferably 10, 15, 20, or 30) contiguous amino acid residues of the amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21

- 35 -

and encompasses an epitope of the protein such that an antibody raised against the peptide forms a specific immune complex with the protein.

Preferred epitopes encompassed by the antigenic peptide are regions of the protein according to the invention that are located on the surface of the protein, e.g., hydrophilic regions, hydrophobic regions, alpha regions, beta regions, coil regions, turn regions and flexible regions.

Immunoassays

Qualitative or quantitative determination of a polypeptide according to the present invention in a biological sample can occur using any art-known method. Antibody-based techniques provide special advantages for assaying specific polypeptide levels in a biological sample.

In these, the specific recognition is provided by the primary antibody (polyclonal or monoclonal) but the secondary detection system can utilize fluorescent, enzyme, or other conjugated secondary antibodies. As a result, an immunocomplex is obtained.

Accordingly, the invention provides a method for diagnosing whether a certain organism is infected with *Aspergillus* comprising the steps of:

- Isolating a biological sample from said organism suspected to be infected with *Aspergillus*,
- reacting said biological sample with an antibody according to the invention,
- determining whether immunocomplexes are formed.

Tissues can also be extracted, e.g., with urea and neutral detergent, for the liberation of protein for Western-blot or dot/slot assay. This technique can also be applied to body fluids.

Other antibody-based methods useful for detecting a protein according to the invention include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). For example, monoclonal antibodies against a protein according to the invention can be used both as an immunoabsorbent and as an enzyme-labeled probe to detect and quantify the protein according to the invention. The amount of protein present in the sample can be calculated by reference to the amount present in a standard preparation using a linear regression computer algorithm. In another ELISA assay, two distinct specific monoclonal antibodies can be used to detect

- 36 -

a protein according to the invention in a biological fluid. In this assay, one of the antibodies is used as the immuno-absorbent and the other as the enzyme-labeled probe.

The above techniques may be conducted essentially as a "one-step" or "two-step" assay. The "one-step" assay involves contacting a protein according to the invention with immobilized antibody and, without washing, contacting the mixture with the labeled antibody. The "two-step" assay involves washing before contacting the mixture with the labeled antibody. Other conventional methods may also be employed as suitable. It is usually desirable to immobilize one component of the assay system on a support, thereby allowing other components of the system to be brought into contact with the component and readily removed from the sample.

Suitable enzyme labels include, for example, those from the oxidase group, which catalyze the production of hydrogen peroxide by reacting with substrate. Activity of an oxidase label may be assayed by measuring the concentration of hydrogen peroxide formed by the enzyme-labelled antibody/substrate reaction.

Besides enzymes, other suitable labels include radioisotopes, such as iodine (^{125}I , ^{127}I), carbon (^{14}C), sulphur (^{35}S), tritium (^3H), indium (^{112}In), and technetium ($^{99\text{m}}\text{Tc}$), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

Specific binding of a test compound to a protein according to the invention can be detected, for example, in vitro by reversibly or irreversibly immobilizing the protein according to the invention on a substrate, e.g., the surface of a well of a 96-well polystyrene microtitre plate. Methods for immobilizing polypeptides and other small molecules are well known in the art. For example, the microtitre plates can be coated with a protein according to the invention by adding the protein in a solution (typically, at a concentration of 0.05 to 1 mg/ml in a volume of 1-100 μl) to each well, and incubating the plates at room temperature to 37 $^{\circ}\text{C}$ for 0.1 to 36 hours. Proteins that are not bound to the plate can be removed by shaking the excess solution from the plate, and then washing the plate (once or repeatedly) with water or a buffer. Typically, the protein is contained in water or a buffer. The plate is then washed with a buffer that lacks the bound protein. To block the free protein-binding sites on the plates, the plates are blocked with a protein that is unrelated to the bound protein. For example, 300 μl of bovine serum albumin (BSA) at a concentration of 2 mg/ml in Tris-HCl is suitable. Suitable substrates include those substrates that contain a defined cross-linking chemistry (e.g., plastic substrates, such as polystyrene, styrene, or polypropylene

- 37 -

substrates from Corning Costar Corp. (Cambridge, MA), for example). If desired, a beaded particle, e.g., beaded agarose or beaded sepharose, can be used as the substrate.

Binding of the test compound to the polypeptides according to the invention can be detected by any of a variety of artknown methods. For example, a specific antibody can be used in an immunoassay. If desired, the antibody can be labeled (e.g., fluorescently or with a radioisotope) and detected directly (see, e.g., West and McMahon, J. Cell Biol. 74:264, 1977). Alternatively, a second antibody can be used for detection (e.g., a labeled antibody that binds the Fc portion of an anti-AN97 antibody). In an alternative detection method, the protein according to the invention is labeled, and the label is detected (e.g., by labeling a protein according to the invention with a radioisotope, fluorophore, chromophore, or the like). In still another method, the protein according to the invention is produced as a fusion protein with a protein that can be detected optically, e.g., green fluorescent protein (which can be detected under UV light). In an alternative method, the protein according to the invention can be covalently attached to or fused with an enzyme having a detectable enzymatic activity, such as horse radish peroxidase, alkaline phosphatase, alpha-galactosidase, or glucose oxidase. Genes encoding all of these enzymes have been cloned and are readily available for use by those of skill in the art. If desired, the fusion protein can include an antigen, and such an antigen can be detected and measured with a polyclonal or monoclonal antibody using conventional methods. Suitable antigens include enzymes (e.g., horse radish peroxidase, alkaline phosphatase, and alpha-galactosidase) and non-enzymatic polypeptides (e.g., serum proteins, such as BSA and globulins, and milk proteins, such as caseins).

Epitopes, antigens and immunogens.

In another aspect, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. These immunogenic epitopes are believed to be confined to a few loci on the molecule. On the other hand, a

- 38 -

region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes. See, for instance, Geysen, H. M. et al., *Proc. Natl. Acad. Sci. USA* 81:3998-4002 (1984).

5 As to the selection of peptides or polypeptides bearing an antigenic epitope (i.e., that contain a region of a protein molecule to which an antibody can bind), it is well known in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein. See, for instance, Sutcliffe, J. G. et al., *Science* 219:660-666 (1984).
10 Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins (i.e., immunogenic epitopes) nor to the amino or carboxyl terminals. Peptides that are extremely hydrophobic and those of six or fewer residues generally are ineffective at
15 inducing antibodies that bind to the mimicked protein; longer, soluble peptides, especially those containing proline residues, usually are effective. Sutcliffe et al., *supra*, at 661. For instance, 18 of 20 peptides designed according to these guidelines, containing 8-39 residues covering 75% of the sequence of the influenza virus hemagglutinin HAI polypeptide chain, induced antibodies that reacted with the HA1
20 protein or intact virus; and 12/12 peptides from the MuLV polymerase and 18/18 from the rabies glycoprotein induced antibodies that precipitated the respective proteins.

 Antigenic epitope-bearing peptides and polypeptides of the invention are therefore useful to raise antibodies, including monoclonal antibodies, that bind specifically to a polypeptide of the invention. Thus, a high proportion of hybridomas
25 obtained by fusion of spleen cells from donors immunized with an antigen epitope-bearing peptide generally secrete antibody reactive with the native protein. Sutcliffe et al., *supra*, at 663. The antibodies raised by antigenic epitope bearing peptides or polypeptides are useful to detect the mimicked protein, and antibodies to different peptides may be used for tracking the fate of various regions of a protein precursor
30 which undergoes posttranslation processing. The peptides and anti-peptide antibodies may be used in a variety of qualitative or quantitative assays for the mimicked protein, for instance in competition assays since it has been shown that even short peptides (e.g., about 9 amino acids) can bind and displace the larger peptides in

- 39 -

immunoprecipitation assays. See, for instance, Wilson, I.A. et al., Cell 37:767-778 at 777 (1984). The anti-peptide antibodies of the invention also are useful for purification of the mimicked protein, for instance, by adsorption chromatography using methods well known in the art.

5 Antigenic epitope-bearing peptides and polypeptides of the invention designed according to the above guidelines preferably contain a sequence of at least seven, more preferably at least nine and most preferably between about 15 to about 30 amino acids contained within the amino acid sequence of a polypeptide of the invention. However, peptides or polypeptides comprising a larger portion of an amino acid sequence of a
10 polypeptide of the invention, containing about 30 to about 50 amino acids, or any length up to and including the entire amino acid sequence of a polypeptide of the invention, also are considered epitope-bearing peptides or polypeptides of the invention and also are useful for inducing antibodies that react with the mimicked protein. Preferably, the amino acid sequence of the epitope-bearing peptide is selected to provide substantial
15 solubility in aqueous solvents (i.e., the sequence includes relatively hydrophilic residues and highly hydrophobic sequences are preferably avoided); and sequences containing proline residues are particularly preferred.

 The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means for making peptides or polypeptides including
20 recombinant means using nucleic acid molecules of the invention. For instance, a short epitope-bearing amino acid sequence may be fused to a larger polypeptide which acts as a carrier during recombinant production and purification, as well as during immunization to produce anti-peptide antibodies.

 Epitope-bearing peptides also may be synthesized using known methods of
25 chemical synthesis. For instance, Houghten has described a simple method for synthesis of large numbers of peptides, such as 10-20 mg of 248 different 13 residue peptides representing single amino acid variants of a segment of the HAI polypeptide which were prepared and characterized (by ELISA-type binding studies) in less than four weeks. Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985). This
30 "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten et al. (1986). In this procedure the individual resins for the solid-phase synthesis of various peptides are contained in separate solvent-permeable packets, enabling the optimal use of the many identical repetitive steps

- 40 -

involved in solid-phase methods.

A completely manual procedure allows 500-1000 or more syntheses to be conducted simultaneously. Houghten et al., *supra*, at 5134.

Epitope-bearing peptides and polypeptides of the invention are used to induce
5 antibodies according to methods well known in the art. See, for instance, Sutcliffe et al.,
supra; Wilson et al., *supra*; Chow, M. et al., *Proc. Natl. Acad. Sci. USA* 82:910-914; and
Bittle, F.J. et al., *J. Gen. Virol.* 66:2347-2354 (1985).

Generally, animals may be immunized with free peptide; however, anti-peptide
antibody titer may be boosted by coupling of the peptide to a macromolecular carrier,
10 such as keyhole limpet hemocyanin (KLH) or tetanus toxoid. For instance, peptides
containing cysteine may be coupled to carrier using a linker such as maleimidobenzoyl-
N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carrier using
a more general linking agent such as glutaraldehyde.

Animals such as rabbits, rats and mice are immunized with either free or
15 carriercoupled peptides, for instance, by intraperitoneal and/or intradermal injection of
emulsions containing about 100 ug peptide or carrier protein and Freund's adjuvant.
Several booster injections may be needed, for instance, at intervals of about two weeks,
to provide a useful titer of anti-peptide antibody which can be detected, for example, by
ELISA assay using free peptide adsorbed to a solid surface. The titer of anti-peptide
20 antibodies in serum from an immunized animal may be increased by selection of anti-
peptide antibodies, for instance, by adsorption to the peptide on a solid support and
elution of the selected antibodies according to methods well known in the art.

Immunogenic epitope-bearing peptides of the invention, i.e., those parts of a
protein that elicit an antibody response when the whole protein is the immunogen, are
25 identified according to methods known in the art. For instance, Geysen et al., 1984,
supra, discloses a procedure for rapid concurrent synthesis on solid supports of
hundreds of peptides of sufficient purity to react in an enzyme-linked immunosorbent
assay. Interaction of synthesized peptides with antibodies is then easily detected without
removing them from the support. In this manner a peptide bearing an immunogenic
30 epitope of a desired protein may be identified routinely by one of ordinary skill in the art.
For instance, the immunologically important epitope in the coat protein of foot-and-mouth
disease virus was located by Geysen et al. with a resolution of seven amino acids by
synthesis of an overlapping set of all 208 possible hexapeptides covering the entire 213

- 41 -

amino acid sequence of the protein. Then, a complete replacement set of peptides in which all 20 amino acids were substituted in turn at every position within the epitope were synthesized, and the particular amino acids conferring specificity for the reaction with antibody were determined. Thus, peptide analogs of the epitope-bearing peptides of the invention can be made routinely by this method. U.S. Patent No. 4,708,781 to Geysen (1987) further describes this method of identifying a peptide bearing an immunogenic epitope of a desired protein.

Further still, U.S. Patent No. 5,194,392 to Geysen (1990) describes a general method of detecting or determining the sequence of monomers (amino acids or other compounds) which is a topological equivalent of the epitope (i.e., a "mimotope") which is complementary to a particular paratope (antigen binding site) of an antibody of interest. More generally, U.S. Patent No. 4,433,092 to Geysen (1989) describes a method of detecting or determining a sequence of monomers which is a topographical equivalent of a ligand which is complementary to the ligand binding site of a particular receptor of interest. Similarly, U.S. Patent No. 5,480,971 to Houghten, R. A. et al. (1996) on Peralkylated Oligopeptide Mixtures discloses linear C1-C7-alkyl peralkylated oligopeptides and sets and libraries of such peptides, as well as methods for using such oligopeptide sets and libraries for determining the sequence of a peralkylated oligopeptide that preferentially binds to an acceptor molecule of interest. Thus, non-peptide analogs of the epitope-bearing peptides of the invention also can be made routinely by these methods.

Use of (hemi)cellulases in industrial processes

The invention also relates to the use of the (hemi)cellulase according to the invention in a selected number of industrial processes. Despite the long term experience obtained with these processes, the (hemi)cellulase according to the invention features a number of significant advantages over the enzymes currently used. Depending on the specific application, these advantages can include aspects like lower production costs, higher specificity towards the substrate, less antigenic, less undesirable side activities, higher yields when produced in a suitable microorganism, more suitable pH and temperature ranges, better tastes of the final product as well as food grade and kosher aspects.

- 42 -

The present invention also relates to methods for preparing a dough or a baked product comprising incorporating into the dough an effective amount of a (hemi)cellulase of the present invention which improves one or more properties of the dough or the baked product obtained from the dough relative to a dough or a baked product in which
5 the polypeptide is not incorporated.

The phrase "incorporating into the dough" is defined herein as adding the (hemi)cellulase according to the invention to the dough, any ingredient from which the dough is to be made, and/or any mixture of dough ingredients from which the dough is to be made. In other words, the (hemi)cellulase according to the invention may be added in
10 any step of the dough preparation and may be added in one, two or more steps. The (hemi)cellulase according to the invention is added to the ingredients of a dough that is kneaded and baked to make the baked product using methods well known in the art. See, for example, U.S. Patent No. 4,567,046, EP-A-426,211, JP-A-60-78529, JP-A-62-111629, and JP-A-63-258528.

15 The term "effective amount" is defined herein as an amount of the (hemi)cellulase according to the invention that is sufficient for providing a measurable effect on at least one property of interest of the dough and/or baked product.

The term "improved property" is defined herein as any property of a dough and/or a product obtained from the dough, particularly a baked product, which is
20 improved by the action of the (hemi)cellulase according to the invention relative to a dough or product in which the (hemi)cellulase according to the invention is not incorporated. The improved property may include, but is not limited to, increased strength of the dough, increased elasticity of the dough, increased stability of the dough, reduced stickiness of the dough, improved extensibility of the dough, improved flavour of
25 the baked product, improved anti-staling of the baked product.

The improved property may be determined by comparison of a dough and/or a baked product prepared with and without addition of a polypeptide of the present invention in accordance with the methods of present invention are described below in the Examples. Organoleptic qualities may be evaluated using procedures well
30 established in the baking industry, and may include, for example, the use of a panel of trained taste-testers.

The term "increased strength of the dough" is defined herein as the property of a dough that has generally more elastic properties and/or requires more work input to

mould and shape.

The term "increased elasticity of the dough" is defined herein as the property of a dough which has a higher tendency to regain its original shape after being subjected to a certain physical strain.

5 The term "increased stability of the dough" is defined herein as the property of a dough that is less susceptible to mechanical abuse thus better maintaining its shape and volume.

10 The term "reduced stickiness of the dough" is defined herein as the property of a dough that has less tendency to adhere to surfaces, e.g., in the dough production machinery, and is either evaluated empirically by the skilled test baker or measured by the use of a texture analyser (e.g., TAXT2) as known in the art.

The term "improved extensibility of the dough" is defined herein as the property of a dough that can be subjected to increased strain or stretching without rupture.

15 The term "improved machineability of the dough" is defined herein as the property of a dough that is generally less sticky and/or more firm and/or more elastic.

The term "increased volume of the baked product" is measured as the specific volume of a given loaf of bread (volume/weight) determined typically by the traditional rapeseed displacement method.

20 The term "improved crumb structure of the baked product" is defined herein as the property of a baked product with finer and/or thinner cell walls in the crumb and/or more uniform/homogenous distribution of cells in the crumb and is usually evaluated empirically by the skilled test baker.

25 The term "improved softness of the baked product" is the opposite of "firmness" and is defined herein as the property of a baked product that is more easily compressed and is evaluated either empirically by the skilled test baker or measured by the use of a texture analyzer (e.g., TAXT2) as known in the art.

The term "improved flavor of the baked product" is evaluated by a trained test panel.

30 The term "improved anti-staling of the baked product" is defined herein as the properties of a baked product that have a reduced rate of deterioration of quality parameters, e.g., softness and/or elasticity, during storage.

The term "dough" is defined herein as a mixture of flour and other ingredients firm enough to knead or roll. The dough may be fresh, frozen, pre-bared, or pre-baked.

- 44 -

The preparation of frozen dough is described by Kulp and Lorenz in Frozen and Refrigerated Doughs and Batters.

The term "baked product" is defined herein as any product prepared from a dough, either of a soft or a crisp character. Examples of baked products, whether of a white, light or dark type, which may be advantageously produced by the present invention are bread (in particular white, whole-meal or rye bread), typically in the form of loaves or rolls, French baguette-type bread, pasta, pita bread, tortillas, tacos, cakes, pancakes, biscuits, cookies, pie crusts, steamed bread, and crisp bread, and the like.

(Hemi)cellulase of the present invention and/or additional enzymes to be used in the methods of the present invention may be in any form suitable for the use in question, e.g., in the form of a dry powder, agglomerated powder, or granulate, in particular a non-dusting granulate, liquid, in particular a stabilized liquid, or protected enzyme such described in WO01/11974 and WO02/26044. Granulates and agglomerated powders may be prepared by conventional methods, e.g., by spraying the (hemi)cellulase according to the invention onto a carrier in a fluid-bed granulator. The carrier may consist of particulate cores having a suitable particle size. The carrier may be soluble or insoluble, e.g., a salt (such as NaCl or sodium sulphate), sugar (such as sucrose or lactose), sugar alcohol (such as sorbitol), starch, rice, corn grits, or soy. The (hemi)cellulase according to the invention and/or additional enzymes may be contained in slow-release formulations. Methods for preparing slow-release formulations are well known in the art. Adding nutritionally acceptable stabilizers such as sugar, sugar alcohol, or another polyol, and/or lactic acid or another organic acid according to established methods may for instance, stabilize liquid enzyme preparations.

The (hemi)cellulase according to the invention may also be incorporated in yeast comprising compositions such as disclosed in EP-A-0619947, EP-A-0659344 and WO02/49441.

For inclusion in pre-mixes of flour it is advantageous that the polypeptide according to the invention is in the form of a dry product, e.g., a non-dusting granulate, whereas for inclusion together with a liquid it is advantageously in a liquid form.

One or more additional enzymes may also be incorporated into the dough. The additional enzyme may be of any origin, including mammalian and plant, and preferably of microbial (bacterial, yeast or fungal) origin and may be obtained by techniques conventionally used in the art.

- 45 -

In a preferred embodiment, the additional enzyme may be an amylase, such as an alpha-amylase (useful for providing sugars fermentable by yeast and retarding staling) or beta-amylase, cyclodextrin glucanotransferase, peptidase, in particular, an exopeptidase (useful in flavour enhancement), transglutaminase, lipase (useful for the
5 modification of lipids present in the dough or dough constituents so as to soften the dough), phospholipase, cellulase, hemicellulase, in particular a pentosanase such as xylanase (useful for the partial hydrolysis of pentosans which increases the extensibility of the dough), protease (useful for gluten weakening in particular when using hard wheat flour), protein disulfide isomerase, e.g., a protein disulfide isomerase as disclosed in WO
10 95/00636, glycosyltransferase, peroxidase (useful for improving the dough consistency), laccase, or oxidase, e.g., an glucose oxidase, hexose oxidase, aldose oxidase, pyranose oxidase, lipoxygenase or L-amino acid oxidase (useful in improving dough consistency).

When one or more additional enzyme activities are to be added in accordance
15 with the methods of the present invention, these activities may be added separately or together with the polypeptide according to the invention, optionally as constituent(s) of the bread-improving and/or dough-improving composition. The other enzyme activities may be any of the enzymes described above and may be dosed in accordance with established baking practices.

20 The present invention also relates to methods for preparing a baked product comprising baking a dough obtained by a method of the present invention to produce a baked product. The baking of the dough to produce a baked product may be performed using methods well known in the art.

25 The present invention also relates to doughs and baked products, respectively, produced by the methods of the present invention.

30 The present invention further relates to a pre-mix, e.g., in the form of a flour composition, for dough and/or baked products made from dough, in which the pre-mix comprises a polypeptide of the present invention. The term "pre-mix" is defined herein to be understood in its conventional meaning, i.e., as a mix of baking agents, generally including flour, which may be used not only in industrial bread-baking plants/facilities, but also in retail bakeries. The pre-mix may be prepared by mixing the polypeptide or a bread-improving and/or dough-improving composition of the invention comprising the polypeptide with a suitable carrier such as flour, starch, a sugar, or a salt. The pre-mix

- 46 -

may contain other dough-improving and/or bread-improving additives, e.g., any of the additives, including enzymes, mentioned above.

The present invention further relates to baking additives in the form of a granulate or agglomerated powder, which comprise a polypeptide of the present invention. The baking additive preferably has a narrow particle size distribution with more
5 than 95% (by weight) of the particles in the range from 25 to 500 μm .

In dough and bread making the present invention may be used in combination with the processing aids defined hereinbefore such as the chemical processing aids like oxidants (e.g. ascorbic acid), reducing agents (e.g. L-cysteine), oxidoreductases (e.g.
10 glucose oxidase) and/or other enzymes such as polysaccharide modifying enzymes (e.g. α -amylase, hemicellulase, branching enzymes, etc.) and/or protein modifying enzymes (endoprotease, exoprotease, branching enzymes, etc.).

EXAMPLE 1

15

Fermentation of *Aspergillus niger*

(Hemi)cellulases encoded by the nucleotide sequence as provided herein were obtained by constructing expression plasmids containing the DNA sequences, transforming an *A. niger* strain with this plasmid and growing the *Aspergillus niger* strains
20 in the following way.

Fresh spores (10^6 - 10^7) of *A. niger* strains were inoculated in 20 ml CSL-medium (100 ml flask, baffle) and grown for 20-24 hours at 34°C and 170 rpm. After inoculation of 5-10 ml CSL pre-culture in 100 ml CSM medium (500 ml flask, baffle) the strains were
25 fermented at 34°C and 170 rpm for 3-5 days.

Cell-free supernatants were obtained by centrifugation in 50 ml Greiner tubes (30 minutes, 5000 rpm). The supernatants were pre-filtered over a GF/A Whatman Glass microfiber filter (150 mm $\varnothing$) to remove the larger particles, adjusted to pH 5 with 4 N KOH (if necessary) and sterile filtrated over a 0.2 μm (bottle-top) filter with suction to
30 remove the fungal material. The supernatants were stored at 4°C (or -20°C).

The CSL medium consisted of (in amount per litre): 100 g Corn Steep Solids (Roquette), 1 g $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$, 0.5 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 10 g glucose $\cdot \text{H}_2\text{O}$ and 0.25 g Basildon (antifoam). The ingredients were dissolved in demi-water and the pH was

- 47 -

adjusted to pH 5.8 with NaOH or H₂SO₄; 100 ml flasks with baffle and foam ball were filled with 20 ml fermentation broth and sterilized for 20 minutes at 120°C after which 200 µl of a solution containing 5000 IU/ml penicillin and 5 mg/ml Streptomycin was added to each flask after cooling to room temperature.

5 The CSM medium consisted of (in amount per litre): 150 g maltose*H₂O, 60 g Soytone (pepton), 1 g NaH₂PO₄*H₂O, 15 g MgSO₄*7H₂O, 0.08 g Tween 80, 0.02 g Basildon (antifoam), 20 g MES, 1 g L-arginine. The ingredients were dissolved in demi-water and the pH was adjusted to pH 6.2 with NaOH or H₂SO₄; 500 ml flasks with baffle and foam ball were filled with 100 ml fermentation broth and sterilized for 20 minutes at
10 120°C after which 1 ml of a solution containing 5000 IU/ml penicillin and 5 mg/ml Streptomycin was added to each flask after cooling to room temperature.

EXAMPLE 2

Purification of the (hemi)cellulases of the invention

15

Step 1 - Preparation of ultrafiltrates

The supernatants of the cultures, as obtained in Example 1, were ultrafiltered to remove the low molecular contaminations that could interfere with the enzymatic activity determinations and the baking tests. Ultrafiltration of 30 ml supernatant was performed in
20 a Millipore Labscale TFF system equipped with a filter with a 10 kDa cut-off.

Depending on their colour, the samples were washed 3–5 times with 40 ml volumes of cold 100 mM phosphate buffer pH 6.0 including 0.5 mM CaCl₂. The final volume of the enzyme solution was 30 ml and is further referred to as "ultrafiltrate".

Step 2 - Determination of the (hemi)cellulases concentration by A280 and HPSEC.

The concentration of the (hemi)cellulases in the ultrafiltrate was calculated from the extinction at 280 nm (A₂₈₀) attributable to the (hemi)cellulases and the calculated molecular extinction coefficient of the (hemi)cellulases. Measurement of the A₂₈₀ was performed in an Uvikon XL Secomam spectrophotometer (Beun de Ronde, Abcoude,
30 The Netherlands):

The molecular extinction coefficient of an enzyme can be calculated from the number of tyrosine, tryptophan and cysteine residues per enzyme molecule (S.C. Gill and P.H. von Hippel, Anal. Biochem. 182, 319-326 (1989)). The molecular extinction

- 48 -

coefficient of these amino acids are 1280, 5690 and 120 $\text{M}^{-1}.\text{cm}^{-1}$ respectively. The number of tyrosine, tryptophan and cysteine residues in the (hemi)cellulases of the invention can be deduced from the protein sequences selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21. The calculated extinction

5 coefficients of the (hemi)cellulases of the invention are summarized in Table 2.

Table 2

(hemi)cellulase	SEQ ID NO:	# amino acids			Calculated M.W. (Da)	Calculated extinction coefficient at 280 nm	
		Trp	Tyr	Cys		$\text{M}^{-1}.\text{cm}^{-1}$	$(1 \text{ mg/ml})^{-1}.\text{cm}^{-1}$
NBE12	3	10	28	3	55928	93100	1.7
NBE21	6	7	16	2	24880	60550	2.4
NBE22	9	8	13	1	27947	62280	2.2
NBE64	12	5	29	8	52508	66530	1.3
NBE75	15	9	13	2	25874	68090	2.6
NBE92	18	10	17	3	27037	79020	2.9
NBE97	21	14	35	12	83328	125900	1.5

The extinction of the ultrafiltrate at 280 nm (A_{280}) that is attributable to the

10 (hemi)cellulases depends on the purity of the enzyme sample. This purity was determined using HPSEC (High Performance Size Exclusion Chromatography) with a TSK SW-XL column (300*7,8 mm; MW range 10-300 kDa). The elution buffer consisted of 25 mM sodium phosphate buffer pH 6.0 and was used at a flow of 1 ml/min. Samples of 5—100 μl were injected. The absorbance at 280 nm was measured.

15 The A_{280} in the ultrafiltrate attributable to the (hemi)cellulase of the invention was obtained from the ratio of the peak surface of the respective (hemi)cellulase peak in the chromatogram and the total surface of the peaks absorbing at 280 nm. The (hemi)cellulase concentration in the ultrafiltrate was then calculated by multiplying the A_{280} of the ultrafiltrate by the ratio described above and divided by the calculated

20 extinction coefficient (1 mg/ml solution – Table 2 most right column) for each (hemi)cellulase.

EXAMPLE 3

Activity measurements

- 49 -

The ultrafiltrates obtained in Example 2, were subjected to the FAU enzyme activity measurement. The activity of the fungal alpha-amylase was measured using Phadebas Amylase test tablets (Pharmacia). Phadebas tablets contain a water insoluble starch substrate and a blue dye, bound by cross-linking to the substrate. The substrate is hydrolised by fungal amylase, releasing dyed soluble maltodextrines that go into solution. A calibration curve was prepared with a solution containing a reference fungal alpha amylase activity.

From the reference and unknown samples appropriate dilutions were prepared in 50 mM malic acid buffer pH 5.5. Samples of 5 ml were incubated with 30°C for 5 minutes, a Phadebas tablet was added and after 15 minutes the reaction was stopped by the addition of 1.0 ml 0.5 N sodium hydroxide. The mixtures were allowed to cool down to room temperature for 5 minutes after which 4.0 ml water was added, shaken by hand and after 15 minutes the samples were centrifuged at 4700 rpm for 10 minutes. The extinction of the top layers was measured at 620 nm. The OD 620 nm is a measure for fungal alpha amylase activity;

One fungal amylase unit (FAU) is defined herein as the amount of enzyme that converts 1 gram of starch (100% dry matter) per hour into a product having a transmission at 620 nm after reaction with a iodine solution of known strength at the reaction conditions stated.

Table 3. FAU activities in the ultrafiltrates as prepared in Example 2.

(Hemi)cellulase	Protein (mg/ml) from the 280 nm analysis	fungal amylase (FAU/ml)
NBE012	2.7	0.12
NBE021	2.5	0
NBE022	0.3	0.55
NBE064	1.7	0.01
NBE075	0.4	2.0
NBE092	1.9	0.01

In addition to the activities mentioned in Table 3, minor activities of glucoamylase was also present, however in such low amounts that these enzymes did

- 50 -

not interfere in the baking experiments described in example 4.

EXAMPLE 4

Baking experiments 1 – pup loaves

5

Pup loaves were baked from 150 gram dough pieces obtained by mixing 200 g flour (Kolibri™/Ibis™ in a ratio of 80/20), 1,4 g dried baker's yeast (Fermipan®), 4 g salt, 3 g sugar, 10 mg ascorbic acid, 116 g water and 2 g fat. After mixing for 6 minutes and 15 seconds in a pin mixer, the dough was divided into pieces of 150 grams and proofed for 45 minutes at 30°C, punched, proofed for another 25 minutes, moulded and panned. Proofing took place at a relative humidity of 90-100%. After a final proof of 70 minutes at 30°C, the dough was baked for 20 minutes at 225°C.

The various effects (Table 4) of the different (hemi)cellulases in the baking experiments were compared with a control containing the same amount of fungal amylase that was added otherwise by the dosage of the ultrafiltrate (for the fungal amylase activity in the ultrafiltrates see Table 3). This was necessary since the amounts of fungal amylase added with the (hemi)cellulases in particular affected the loaf volume, not the other parameters. The volume of the breads with the control amount of fungal amylase added was taken as 100%.

20

Table 4.

effect		Score				
		1	2	3	4	5
dough	dough stickiness	too sticky	sticky	control bread	much better	excellent dry
	dough extensibility	Too short	Shorter than the control	control bread	Good	too long
baked bread	crumb structure	poor	non-uniform	control bread	Good	excellent
	crust colour	Nearly white	too light	control bread	excellent	too dark

- 51 -

	crumb colour	Far too yellow	too yellow	control bread	excellent	absolutely white
--	--------------	----------------	------------	---------------	-----------	------------------

Loaf volume was determined by the Bread Volume Measurer BVM-3 (RI Cards Instruments AB, Viken, Sweden). The principle of this measurement is based on the reflection of ultrasound measured by a sensor around a rotating bread. A measurement time was taken of 45 seconds.

Dough stickiness and extensibility were evaluated by a qualified baker using the scale depicted in Table 4. The average of 2 loaves per object was measured.

After these tests the dough pieces were rounded and a first proof was performed for 45 minutes at 30°C and hereafter the dough was punched, moulded, panned, proofed for 75 minutes at 30°C. The relative humidity during the proofs was set at 85%.

Subsequently the stability of the proofed dough was judged by the presence of bladders, torn side crust and irregular curved surfaces of the crust. The dough pieces were baked for 20 minutes at 225°C. Loaf volumes were determined by the BVM-3 method: in the table the average is presented of 2 breads that are baked from the same object.

The crumb structure was judged by a qualified baker using the scale depicted in Table 4. After storing the loaves for three days in polyethylene bags at room temperature crumb firmness was measured using a Stevens Texture Analyser. Two slices of 2 cm thickness from the centre of each loaf were analysed by the texture analyser using a probe of 1.5 inch diameter, a compression depth of 5 mm (25%) and a rate of compression of 0.5 mm/sec. In the table the average is shown of two measurements.

Crust colour was judged by a qualified baker according to the scale depicted in Table 4. As a reference the standard recipe for Dutch tin bread was used.

Crumb colour was judged by a qualified baker according to the scale depicted in Table 4. The colour of the crumb of the control breads was judged as normal (3). As a positive control the breads of 2 objects are used with the same composition as the control plus 0.5% soya flour. The proofing and baking procedure are the same as that of the control without soya flour. The latter is judged as "excellent".

The overhanging top of the bread was judged by the hanging of the top in

- 52 -

relation to the baking tin, the lower the edges of the top the lower the judgement. The less hanging, the better the judgement,

Staling of the bread was judged by feeling the firmness of the crumb of slices of the bread. Before slicing took place, the bread was stored in a plastic bag at room temperature for 4 days. The softer the crumb of the slices is, the better the judgement.

Table 5. Baking performance of the (hemi)cellulases of the invention

parameter	(hemi)cellulases					
	NBE012	NBE021	NBE022	NBE064	NBE075	NBE092
Volume (%)	113	107	108	112	109	107
dough stickiness	3	3	2	4	3	2
dough extensibility	2	3	4	3	2	4
dough stability	4	4	4	4	4	4
crumb structure	4	4	4	4	3	4
crust colour	3	3	4	4	4	3
crumb colour	4	4	4	4	4	3
overhanging top	3	3	3	3	4	3
staling	4	not tested	3	3	3	4

- 53 -

CLAIMS

1. An isolated polynucleotide hybridisable to a polynucleotide selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20.
5
2. An isolated polynucleotide according to claim 1 hybridisable under high stringency conditions to a polynucleotide selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20.
3. An isolated polynucleotide according to claims 1 or 2 obtainable from a filamentous fungus.
10
4. An isolated polynucleotide according to claim 3 obtainable from *Aspergillus niger*.
5. An isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or functional equivalents thereof.
15
6. An isolated polynucleotide encoding at least one functional domain of a polypeptide selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or functional equivalents thereof.
7. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20 or functional equivalents thereof.
20
8. An isolated polynucleotide selected from the group consisting of SEQ ID NO: 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19 and 20.
9. A vector comprising a polynucleotide sequence according to claims 1 to 8.
- 25 10. A vector according to claim 9 wherein said polynucleotide sequence according to claims 1 to 8 is operatively linked with regulatory sequences suitable for expression of said polynucleotide sequence in a suitable host cell.

- 54 -

11. A vector according to claim 10 wherein said suitable host cell is a filamentous fungus.
12. A method for manufacturing a polynucleotide according to claims 1 – 8 or a vector according to claims 9 to 11 comprising the steps of culturing a host cell transformed with said polynucleotide or said vector and isolating said polynucleotide or said vector from said host cell.
13. An isolated (hemi)cellulase selected from the group consisting of SEQ ID NO: 3, 6, 9, 12, 15, 18 and 21 or functional equivalents thereof.
14. An isolated (hemi)cellulase according to claim 13 obtainable from *Aspergillus niger*.
15. An isolated (hemi)cellulase obtainable by expressing a polynucleotide according to claims 1 to 8 or a vector according to claims 9 to 11 in an appropriate host cell, e.g. *Aspergillus niger*.
16. Recombinant (hemi)cellulase comprising a functional domain of any of the (hemi)cellulases according to claims 13-15.
17. A method for manufacturing a (hemi)cellulase according to claims 13 to 16 comprising the steps of transforming a suitable host cell with an isolated polynucleotide according to claims 1 to 8 or a vector according to claims 9 to 11, culturing said cell under conditions allowing expression of said polynucleotide and optionally purifying the encoded polypeptide from said cell or culture medium.
18. A recombinant host cell comprising a polynucleotide according to claims 1 to 8 or a vector according to claims 9 to 11.
19. A recombinant host cell expressing a (hemi)cellulase according to claims 13 to 16.
20. Purified antibodies reactive with a (hemi)cellulase according to claims 13 to 16.
21. Fusion protein comprising a (hemi)cellulase sequence according to claims 13 to 16.

- 55 -

- 22. A process for the production of dough comprising adding a (hemi)cellulase according to anyone of claims 13-16.
- 23. A process for the production of a baked product from a dough as prepared by the process of claim 22.
- 5 24. Use of a (hemi)cellulase according to anyone of claims 13-16 for the preparation of a dough and/or the baked product thereof.

21062WO.ST25.txt
SEQUENCE LISTING

<110> DSM IP Assets B.V.

<120> NOVEL (HEMI)CELLULASES AND USES THEREOF

<130> 21062WO

<160> 21

<170> PatentIn version 3.1

<210> 1

<211> 2596

<212> DNA

<213> *Aspergillus niger*

<400> 1

cgcgatgtga	ctgtggcacc	tcgccaagtc	tcaagttaaa	ttaagcgccc	accctgaggt	60
atctgatggg	tatgaagtgt	ggacggagta	cctgctgggc	tgccctgcacg	gccagccagg	120
aacaagtacc	gatcacatca	ggaatacccc	tcggggacca	gcagaggaat	ccacccttgc	180
atggcttgcc	accaatcgga	ggcccagatc	ggcgggagtg	gacctgaagc	agcttagcgg	240
agtggggctt	gatcccgccc	aggaaccagg	gactgctggc	tgccctgcct	ctttgggaaa	300
ctccagggcc	aggttgggtt	gctggagggc	gcttccatct	tcccctgtgc	cgtcattgta	360
tgcaaagcag	ttctgtgtgt	cctcgatgca	atactgcact	aaacgagtag	tctgcgggtc	420
tccccaaac	tggatgcagc	ataactctcc	ttctcagcct	gatcccgctc	gcttctccgc	480
attcttttagc	cggtagatct	tcttatttat	acccgattcc	gtccgctcct	tgacgtgctc	540
tcattcttgc	gtcctgtttc	cgtcgtttgt	gcaagacgag	cttggttttt	ctattcgggt	600
ctgctcactc	ggttgctact	catttcgcca	aaccatgcat	ctcagctcgc	tttccttgtc	660
cctgacggcc	ttggccatcg	tcagtccttc	tgtctcttat	cctcacttgg	gttcttccca	720
gcctgtgctt	cacactaatt	ccgacaccac	tcaatcgcg	gcagacgcca	tcaaggctgc	780
cttctcgac	gcctgggatg	gctacctcca	atacgcttc	cctcatgacg	agctgcaccc	840
cgtctcgaat	ggctatggcg	actcccggaa	tggatggggc	gcctctgccc	tggatgcact	900
ctccacagca	gtgattatgc	gcaatgctac	cattgtcaat	cagatcctgg	accacgtggc	960
caagattgac	tactccaaga	caaacaccac	cgtgagcctc	ttcgagacca	ccatccgcta	1020
cctgggtggc	atgctgtccg	gatatgacct	gctgaagggg	cccgtgtccg	acctggtgca	1080
agactcgtcg	aaaatcgacg	tgctcttgac	ccaatcgaag	aacctcggcg	acgttctcaa	1140
attcgccctc	gatacacctt	cgggggttcc	atacaacaat	ctcaacatta	cctcaggggg	1200
gaatgatgga	gccaagacca	atggctcttg	tgtgacgggt	acactggcgc	tggaatggac	1260
gcgcctgtcg	gaccttacgg	gtgacaccac	gtatgccgac	ctgagccaga	aggccgagtc	1320
ttacctctc	aacccgcagc	ctaagtctgc	ggagccattc	ccgggcttgg	taggaagcaa	1380
catcaacatc	agtaatgggc	agttcacgga	tgcgcagggt	agctggaatg	gcggcgatga	1440
ctcttactac	gagtacctga	tcaaaatgta	cgtatacgac	cccaagcggg	tcggtctcta	1500

21062W0.ST25.txt

taaagaccgc tgggtggccg cggcgcaatc gaccatgcag catctggcgt cgcacccctc	1560
gacccgcccc gacctgacct tcctggcctc ctacaacaac ggaacttttg ggttgagctc	1620
gcagcacctg acttgcttcg acggaggtag cttcctgttg ggaggcaccg tgctgaaccg	1680
gacggatttc attaatctcg gcctggacct agtgagcggg tgtcatgata catacaactc	1740
gacgttgacg ggaattggcc cggagtcgtt cagctgggat accagcgaca tcccctcgag	1800
ccagcagtcg ctctacgaga aagccggatt ctacattact agtggcgcgt acattttgcg	1860
cccgaagtg attgagagtt tctactatgc ctggagagct accgggcagg aaacggtaag	1920
ttcgacttct cccccttttg ctcaatagta actaacaacc atatcagtac cggaatgga	1980
tctggagcgc attcagcgca gtcaacgact actgccgcac cgactctgga ttttcaggctc	2040
tgactgacgt gaacgcggcc aacggcggca gccggtacga caaccaggag agtttcctgt	2100
ttgccgaggt gatgaagtat tcgtacatgg cgttcgcgga agatgcagcc tggcaggtgc	2160
agcccggcag cggaaccaa tttgtgttca acaccgaggc acaccagtg cgggtgagta	2220
gtacataaat ggtagacatt gttagagatg gttttggtag caatagagga cagctcaagt	2280
cacctaaata ctctgcagat ctgtttctgc cttgttttaa atataccgta gacaagaaac	2340
ccccgcaaac tcgtcatac tccctcccat tctttgttcc attcagtcga tcctcagaag	2400
atgataggcg gtccaccgcc ccacccccat acaaaatagt tctgtggcgc cccacaccct	2460
gagatgcgca catttgagca tctttgcttt cagtgtacc ttgggggaac atgccgatat	2520
catgctatgt atgtgtttac gtaccttggg ggctgggcaa tatagacata gtagtcagct	2580
gtagccccta cagcct	2596

<210> 2

<211> 1542

<212> DNA

<213> *Aspergillus niger*

<220>

<221> CDS

<222> (1)..(1542)

<400> 2

atg cat ctc agc tcg ctt tcc ttg tcc ctg acg gcc ttg gcc atc gtc	48
Met His Leu Ser Ser Leu Ser Leu Ser Leu Thr Ala Leu Ala Ile Val	
1 5 10 15	
agt cct tct gct gct tat cct cac ttg ggt tct tcc cag cct gtg ctt	96
Ser Pro Ser Ala Ala Tyr Pro His Leu Gly Ser Ser Gln Pro Val Leu	
20 25 30	
cac act aat tcc gac acc act caa tcg cgc gca gac gcc atc aag gct	144
His Thr Asn Ser Asp Thr Thr Gln Ser Arg Ala Asp Ala Ile Lys Ala	
35 40 45	
gcc ttc tcg cac gcc tgg gat ggc tac ctc caa tac gcc ttc cct cat	192
Ala Phe Ser His Ala Trp Asp Gly Tyr Leu Gln Tyr Ala Phe Pro His	
50 55 60	
gac gag ctg cac ccc gtc tcg aat ggc tat ggc gac tcc cgg aat gga	240

21062WO.ST25.txt

Asp	Glu	Leu	His	Pro	Val	Ser	Asn	Gly	Tyr	Gly	Asp	Ser	Arg	Asn	Gly		
65					70					75					80		
tgg	ggc	gcc	tct	gcc	gtg	gat	gca	ctc	tcc	aca	gca	gtg	att	atg	cgc		288
Trp	Gly	Ala	Ser	Ala	Val	Asp	Ala	Leu	Ser	Thr	Ala	Val	Ile	Met	Arg		
				85					90					95			
aat	gct	acc	att	gtc	aat	cag	atc	ctg	gac	cac	gtg	gcc	aag	att	gac		336
Asn	Ala	Thr	Ile	Val	Asn	Gln	Ile	Leu	Asp	His	Val	Ala	Lys	Ile	Asp		
				100					105					110			
tac	tcc	aag	aca	aac	acc	acc	gtg	agc	ctc	ttc	gag	acc	acc	atc	cgc		384
Tyr	Ser	Lys	Thr	Asn	Thr	Thr	Val	Ser	Leu	Phe	Glu	Thr	Thr	Ile	Arg		
				115					120					125			
tac	ctg	ggt	ggc	atg	ctg	tcc	gga	tat	gac	ctg	ctg	aag	ggg	ccc	gtg		432
Tyr	Leu	Gly	Gly	Met	Leu	Ser	Gly	Tyr	Asp	Leu	Leu	Lys	Gly	Pro	Val		
				130					135					140			
tcc	gac	ctg	gtg	caa	gac	tcg	tcg	aaa	atc	gac	gtg	ctc	ttg	acc	caa		480
Ser	Asp	Leu	Val	Gln	Asp	Ser	Ser	Lys	Ile	Asp	Val	Leu	Leu	Thr	Gln		
				145					150					155		160	
tcg	aag	aac	ctc	ggc	gac	gtt	ctc	aaa	ttc	gcc	ttc	gat	aca	ccc	tcg		528
Ser	Lys	Asn	Leu	Gly	Asp	Val	Leu	Lys	Phe	Ala	Phe	Asp	Thr	Pro	Ser		
				165					170					175			
ggg	gtt	cca	tac	aac	aat	ctc	aac	att	acc	tca	ggg	ggg	aat	gat	gga		576
Gly	Val	Pro	Tyr	Asn	Asn	Leu	Asn	Ile	Thr	Ser	Gly	Gly	Asn	Asp	Gly		
				180					185					190			
gcc	aag	acc	aat	ggt	ctt	gct	gtg	acg	ggt	aca	ctg	gcg	ctg	gaa	tgg		624
Ala	Lys	Thr	Asn	Gly	Leu	Ala	Val	Thr	Gly	Thr	Leu	Ala	Leu	Glu	Trp		
				195					200					205			
acg	cgc	ctg	tcg	gac	ctt	acg	ggt	gac	acc	acg	tat	gcc	gac	ctg	agc		672
Thr	Arg	Leu	Ser	Asp	Leu	Thr	Gly	Asp	Thr	Thr	Tyr	Ala	Asp	Leu	Ser		
				210					215					220			
cag	aag	gcc	gag	tct	tac	ctc	ctc	aac	ccg	cag	cct	aag	tct	gcg	gag		720
Gln	Lys	Ala	Glu	Ser	Tyr	Leu	Leu	Asn	Pro	Gln	Pro	Lys	Ser	Ala	Glu		
				225					230					235		240	
cca	ttc	ccg	ggc	ttg	gta	gga	agc	aac	atc	aac	atc	agt	aat	ggg	cag		768
Pro	Phe	Pro	Gly	Leu	Val	Gly	Ser	Asn	Ile	Asn	Ile	Ser	Asn	Gly	Gln		
				245					250					255			
ttc	acg	gat	gcg	cag	gtg	agc	tgg	aat	ggc	ggc	gat	gac	tct	tac	tac		816
Phe	Thr	Asp	Ala	Gln	Val	Ser	Trp	Asn	Gly	Gly	Asp	Asp	Ser	Tyr	Tyr		
				260					265					270			
gag	tac	ctg	atc	aaa	atg	tac	gta	tac	gac	ccc	aag	cgg	ttc	ggt	ctc		864
Glu	Tyr	Leu	Ile	Lys	Met	Tyr	Val	Tyr	Asp	Pro	Lys	Arg	Phe	Gly	Leu		
				275					280					285			
tat	aaa	gac	cgc	tgg	gtg	gcc	gcg	gcg	caa	tcg	acc	atg	cag	cat	ctg		912
Tyr	Lys	Asp	Arg	Trp	Val	Ala	Ala	Ala	Gln	Ser	Thr	Met	Gln	His	Leu		
				290					295					300			

gcg	tcg	cac	ccc	tcg	acc	cgc	ccg	gac	ctg	acc	ttc	ctg	gcc	tcc	tac		960
Ala	Ser	His	Pro	Ser	Thr	Arg	Pro	Asp	Leu	Thr	Phe	Leu	Ala	Ser	Tyr		
305					310					315					320		
aac	aac	gga	act	ttg	ggg	ttg	agc	tcg	cag	cac	ctg	act	tgc	ttc	gac		1008
Asn	Asn	Gly	Thr	Leu	Gly	Leu	Ser	Ser	Gln	His	Leu	Thr	Cys	Phe	Asp		
				325					330					335			
gga	ggt	agc	ttc	ctg	ttg	gga	ggc	acc	gtg	ctg	aac	cgg	acg	gat	ttc		1056
Gly	Gly	Ser	Phe	Leu	Leu	Gly	Gly	Thr	Val	Leu	Asn	Arg	Thr	Asp	Phe		
			340					345					350				
att	aat	ttc	ggc	ctg	gac	cta	gtg	agc	ggg	tgt	cat	gat	aca	tac	aac		1104
Ile	Asn	Phe	Gly	Leu	Asp	Leu	Val	Ser	Gly	Cys	His	Asp	Thr	Tyr	Asn		
		355					360					365					
tcg	acg	ttg	acg	gga	att	ggc	ccg	gag	tcg	ttc	agc	tgg	gat	acc	agc		1152
Ser	Thr	Leu	Thr	Gly	Ile	Gly	Pro	Glu	Ser	Phe	Ser	Trp	Asp	Thr	Ser		
	370					375					380						
gac	atc	ccc	tcg	agc	cag	cag	tcg	ctc	tac	gag	aaa	gcc	gga	ttc	tac		1200
Asp	Ile	Pro	Ser	Ser	Gln	Gln	Ser	Leu	Tyr	Glu	Lys	Ala	Gly	Phe	Tyr		
385					390					395					400		
att	act	agt	ggc	gcg	tac	att	ttg	cgc	ccg	gaa	gtg	att	gag	agt	ttc		1248
Ile	Thr	Ser	Gly	Ala	Tyr	Ile	Leu	Arg	Pro	Glu	Val	Ile	Glu	Ser	Phe		
			405						410					415			
tac	tat	gcc	tgg	aga	gct	acc	ggg	cag	gaa	acg	tac	cgg	gaa	tgg	atc		1296
Tyr	Tyr	Ala	Trp	Arg	Ala	Thr	Gly	Gln	Glu	Thr	Tyr	Arg	Glu	Trp	Ile		
		420						425					430				
tgg	agc	gca	ttc	agc	gca	gtc	aac	gac	tac	tgc	cgc	acc	gac	tct	gga		1344
Trp	Ser	Ala	Phe	Ser	Ala	Val	Asn	Asp	Tyr	Cys	Arg	Thr	Asp	Ser	Gly		
		435					440					445					
ttt	tca	ggt	ctg	act	gac	gtg	aac	gcg	gcc	aac	ggc	ggc	agc	cgg	tac		1392
Phe	Ser	Gly	Leu	Thr	Asp	Val	Asn	Ala	Ala	Asn	Gly	Gly	Ser	Arg	Tyr		
	450					455				460							
gac	aac	cag	gag	agt	ttc	ctg	ttt	gcc	gag	gtg	atg	aag	tat	tcg	tac		1440
Asp	Asn	Gln	Glu	Ser	Phe	Leu	Phe	Ala	Glu	Val	Met	Lys	Tyr	Ser	Tyr		
465					470				475						480		
atg	gcg	ttc	gcg	gaa	gat	gca	gcc	tgg	cag	gtg	cag	ccc	ggc	agc	ggg		1488
Met	Ala	Phe	Ala	Glu	Asp	Ala	Ala	Trp	Gln	Val	Gln	Pro	Gly	Ser	Gly		
			485						490					495			
aac	caa	ttt	gtg	ttc	aac	acc	gag	gca	cac	cca	gtg	cgg	gtg	agt	agt		1536
Asn	Gln	Phe	Val	Phe	Asn	Thr	Glu	Ala	His	Pro	Val	Arg	Val				

21062WO.ST25.txt

<211> 513

<212> PRT

<213> *Aspergillus niger*

<400> 3

```

Met His Leu Ser Ser Leu Ser Leu Ser Leu Thr Ala Leu Ala Ile Val
1      5      10      15
Ser Pro Ser Ala Ala Tyr Pro His Leu Gly Ser Ser Gln Pro Val Leu
      20      25      30
His Thr Asn Ser Asp Thr Thr Gln Ser Arg Ala Asp Ala Ile Lys Ala
      35      40      45
Ala Phe Ser His Ala Trp Asp Gly Tyr Leu Gln Tyr Ala Phe Pro His
      50      55      60
Asp Glu Leu His Pro Val Ser Asn Gly Tyr Gly Asp Ser Arg Asn Gly
65      70      75      80
Trp Gly Ala Ser Ala Val Asp Ala Leu Ser Thr Ala Val Ile Met Arg
      85      90      95
Asn Ala Thr Ile Val Asn Gln Ile Leu Asp His Val Ala Lys Ile Asp
      100     105     110
Tyr Ser Lys Thr Asn Thr Thr Val Ser Leu Phe Glu Thr Thr Ile Arg
      115     120     125
Tyr Leu Gly Gly Met Leu Ser Gly Tyr Asp Leu Leu Lys Gly Pro Val
      130     135     140
Ser Asp Leu Val Gln Asp Ser Ser Lys Ile Asp Val Leu Leu Thr Gln
145     150     155     160
Ser Lys Asn Leu Gly Asp Val Leu Lys Phe Ala Phe Asp Thr Pro Ser
      165     170     175
Gly Val Pro Tyr Asn Asn Leu Asn Ile Thr Ser Gly Gly Asn Asp Gly
      180     185     190
Ala Lys Thr Asn Gly Leu Ala Val Thr Gly Thr Leu Ala Leu Glu Trp
      195     200     205
Thr Arg Leu Ser Asp Leu Thr Gly Asp Thr Thr Tyr Ala Asp Leu Ser
      210     215     220
Gln Lys Ala Glu Ser Tyr Leu Leu Asn Pro Gln Pro Lys Ser Ala Glu
225     230     235     240
Pro Phe Pro Gly Leu Val Gly Ser Asn Ile Asn Ile Ser Asn Gly Gln
      245     250     255
Phe Thr Asp Ala Gln Val Ser Trp Asn Gly Gly Asp Asp Ser Tyr Tyr
      260     265     270
Glu Tyr Leu Ile Lys Met Tyr Val Tyr Asp Pro Lys Arg Phe Gly Leu
      275     280     285
Tyr Lys Asp Arg Trp Val Ala Ala Ala Gln Ser Thr Met Gln His Leu
      290     295     300
Ala Ser His Pro Ser Thr Arg Pro Asp Leu Thr Phe Leu Ala Ser Tyr

```

21062WO.ST25.txt

```

305          310          315          320
Asn Asn Gly Thr Leu Gly Leu Ser Ser Gln His Leu Thr Cys Phe Asp
          325          330          335
Gly Gly Ser Phe Leu Leu Gly Gly Thr Val Leu Asn Arg Thr Asp Phe
          340          345          350
Ile Asn Phe Gly Leu Asp Leu Val Ser Gly Cys His Asp Thr Tyr Asn
          355          360          365
Ser Thr Leu Thr Gly Ile Gly Pro Glu Ser Phe Ser Trp Asp Thr Ser
          370          375          380
Asp Ile Pro Ser Ser Gln Gln Ser Leu Tyr Glu Lys Ala Gly Phe Tyr
385          390          395          400
Ile Thr Ser Gly Ala Tyr Ile Leu Arg Pro Glu Val Ile Glu Ser Phe
          405          410          415
Tyr Tyr Ala Trp Arg Ala Thr Gly Gln Glu Thr Tyr Arg Glu Trp Ile
          420          425          430
Trp Ser Ala Phe Ser Ala Val Asn Asp Tyr Cys Arg Thr Asp Ser Gly
          435          440          445
Phe Ser Gly Leu Thr Asp Val Asn Ala Ala Asn Gly Gly Ser Arg Tyr
          450          455          460
Asp Asn Gln Glu Ser Phe Leu Phe Ala Glu Val Met Lys Tyr Ser Tyr
465          470          475          480
Met Ala Phe Ala Glu Asp Ala Ala Trp Gln Val Gln Pro Gly Ser Gly
          485          490          495
Asn Gln Phe Val Phe Asn Thr Glu Ala His Pro Val Arg Val Ser Ser
          500          505          510
Thr

```

<210> 4

<211> 2439

<212> DNA

<213> *Aspergillus niger*

<400> 4

```

agagtctgag tttaaaatgc agagtctcgc tggaagatac taacctgaac gccaatctag      60
cttcctcctg ttttccagtc gccgtggggc aaagcatatt agctgacagg atggcgctcg      120
atgtgaggaa gtagcggatg cttaaaatag ccaccactac tactattatt tgttgggagg      180
agactcagat cgtgacgttc ctgtcaatat tttagaaata ggggcccaaa agtatcgaga      240
cgaaaggcat gcatatgtat catggaaacc catcatcagc gattccattg atggattcaa      300
ccgtggcgag aactacaaga tacagtacgc tattctgaga tgtgcctcta ctcagtgcct      360
ctccctagca ataagaattc ggccaaaaag tcccgtcccc ttaggaaggg gcactgcaat      420
atgccctg acggatgctg cagcactat caacacacta atcggtagga gcatccgact      480
ggaagctcca ggggagcagc tactttgcc cagtggactt gtatgccatt agggagtctt      540
agaccacga ctgatacaga actcgactcc ttggtcatgc ccaatagact gtttcgtcaa      600

```

21062W0.ST25.txt

gactaacaga	tagtagtata	ggcctttatc	acatcgacat	gatggtgcc	ttgcatcgac	660
gataggcgaa	tgacaccata	ttcaatcaat	gatatgaaga	gtcgtcatgc	gacaactaag	720
gatcatcttt	caccaatgct	gtcctcctg	gcagacttgt	acaaatcctc	tcggctgcat	780
taacggcggg	cccatTTTTA	acgggactct	tggtgagcct	ttttctttgc	ataacttagc	840
tttgaccacc	caggggcttc	agcattcagt	tcgtgaagaa	caagaaaggc	gctaaaggga	900
tattattgct	aaaaaaggat	tccattagca	gggccggtgg	aggcgcgtcg	gtgtgccaga	960
gggagaccaa	cgataccggc	aatctcggcg	gatcttcttg	ctggggaaaa	cggccgagtc	1020
caccttgac	accgggcca	attattcctg	ctatcagcac	cattcattgt	gtcttcatgg	1080
ctggatgggc	gttggtgagt	ggtcgatgct	gagggctggc	tcggtcggag	cgccaaagag	1140
taaaagtggg	cgggagtggc	acccactttt	aatttgaaac	accacggcct	tgaataagat	1200
tgtccaggaa	caccaccagg	ggcccaggac	agccgattta	atcaatgaca	gaacgatgag	1260
tgctcgacta	ccctgctcgg	aagcaagctc	gattggatgc	tagatggagc	atcgtaccag	1320
agcacagctg	ctgttcacct	tcatataagg	acggtcgcgg	aacctgggtc	ctgcacagca	1380
acctacgctc	ttttcccgaa	tatttccggc	gcctgctaaa	taattattgt	cactccctta	1440
aactacttca	tcatggtcgc	ctactcgtct	cttctcctcg	ctctttcggc	tgcctcgagc	1500
gccgtcgctc	tccccaatgg	caaggccctg	ctggacatgg	ccaagcgag	tcttacttct	1560
tccgcggaag	gcgagagcaa	cggctacttc	tactccttct	ggaccaacgg	tgagggcgac	1620
gtgacctaca	ccaacggcga	cgctggtgaa	tacaccgttg	agtggaccga	ctgtggtgac	1680
tttgtggccg	gcaagggttg	gaaccggggc	agtgtcagct	aaggattttt	ccatcgctgt	1740
ttatcgaata	aagatgggtg	ttaaccggta	tattcagggc	cgttacctac	agcggatcat	1800
gggaaaccga	cgccaacgcc	tacctgtccg	tctacggttg	gaccacctcc	cccttggtcg	1860
aattttatat	cgtcgacaag	tacggtgatt	acgaccctc	ctctggcctg	accgagctcg	1920
gcaccgtcga	cagcgacgac	ggaacctaca	agatctacca	gaccacctcg	gaggacgccg	1980
actcgatcga	aggcactgcc	accttcaagc	agtactggtc	cgtgcgcact	gagggccgta	2040
ttggtggaac	cgtcaccacg	cagaaccact	tcgatgcctg	ggagaacctt	ggtctggagc	2100
tgggtacctt	caactacatg	atcgttgcc	ccgagggata	cgagagcagc	gggtctgcca	2160
ccatcaccgt	tgagacttct	tctgacagca	ccgaggagag	ctgctaagcg	ctcttttgaa	2220
gtgactgctt	tactccttct	gtatataatc	agtcccatc	gggtttcact	tcttgaccga	2280
actcataatc	tcattgtcac	atctgggtgg	ggacgactag	gccgcgacat	tgttcttcaa	2340
catttgtttt	cgtattatct	taccgcgtat	aggcattcat	agattacacc	ttcctttag	2400
cggctgtaac	cgtcgcagtg	ggagtgcac	gagcagcag			2439

<210> 5

<211> 696

<212> DNA

<213> Aspergillus niger

<220>

<221> CDS

<222> (1)..(696)

<400> 5

atg gtc gcc tac tcg tct ctt ctc ctc gct ctt tcg gct gcc tcg agc

48

21062W0.ST25.txt

Met Val Ala Tyr Ser Ser Leu Leu Leu Ala Leu Ser Ala Ala Ser Ser	
1 5 10 15	
gcc gtc gct ctc ccc aat ggc aag gcc ctg ctg gac atg gcc aag cgc	96
Ala Val Ala Leu Pro Asn Gly Lys Ala Leu Leu Asp Met Ala Lys Arg	
20 25 30	
agt ctt act tct tcc gcg gaa ggc gag agc aac ggc tac ttc tac tcc	144
Ser Leu Thr Ser Ser Ala Glu Gly Glu Ser Asn Gly Tyr Phe Tyr Ser	
35 40 45	
ttc tgg acc aac ggt gag ggc gac gtg acc tac acc aac ggc gac gct	192
Phe Trp Thr Asn Gly Glu Gly Asp Val Thr Tyr Thr Asn Gly Asp Ala	
50 55 60	
ggt gaa tac acc gtt gag tgg acc gac tgt ggt gac ttt gtg gcc ggc	240
Gly Glu Tyr Thr Val Glu Trp Thr Asp Cys Gly Asp Phe Val Ala Gly	
65 70 75 80	
aag ggt tgg aac ccg ggc agt gct cag gcc gtt acc tac agc gga tca	288
Lys Gly Trp Asn Pro Gly Ser Ala Gln Ala Val Thr Tyr Ser Gly Ser	
85 90 95	
tgg gaa acc gac gcc aac gcc tac ctg tcc gtc tac ggt tgg acc acc	336
Trp Glu Thr Asp Ala Asn Ala Tyr Leu Ser Val Tyr Gly Trp Thr Thr	
100 105 110	
tcc ccc ttg gtc gaa ttt tat atc gtc gac aag tac ggt gat tac gac	384
Ser Pro Leu Val Glu Phe Tyr Ile Val Asp Lys Tyr Gly Asp Tyr Asp	
115 120 125	
ccc tcc tct ggc ctg acc gag ctc ggc acc gtc gac agc gac gac gga	432
Pro Ser Ser Gly Leu Thr Glu Leu Gly Thr Val Asp Ser Asp Asp Gly	
130 135 140	
acc tac aag atc tac cag acc acc cgt gag gac gcc gac tcg atc gaa	480
Thr Tyr Lys Ile Tyr Gln Thr Thr Arg Glu Asp Ala Asp Ser Ile Glu	
145 150 155 160	
ggc act gcc acc ttc aag cag tac tgg tcc gtg cgc act gag ggc cgt	528
Gly Thr Ala Thr Phe Lys Gln Tyr Trp Ser Val Arg Thr Glu Gly Arg	
165 170 175	
att ggt gga acc gtc acc acg cag aac cac ttc gat gcc tgg gag aac	576
Ile Gly Gly Thr Val Thr Thr Gln Asn His Phe Asp Ala Trp Glu Asn	
180 185 190	
ctt ggt ctg gag ctg ggt acc ttc aac tac atg atc gtt gcc acc gag	624
Leu Gly Leu Glu Leu Gly Thr Phe Asn Tyr Met Ile Val Ala Thr Glu	
195 200 205	
gga tac gag agc agc ggg tct gcc acc atc acc gtt gag act tct tct	672
Gly Tyr Glu Ser Ser Gly Ser Ala Thr Ile Thr Val Glu Thr Ser Ser	
210 215 220	
gac agc acc gag gag agc tgc taa	696
Asp Ser Thr Glu Glu Ser Cys	
225 230	

21062WO.ST25.txt

<210> 6

<211> 231

<212> PRT

<213> Aspergillus niger

<400> 6

```

Met Val Ala Tyr Ser Ser Leu Leu Leu Ala Leu Ser Ala Ala Ser Ser
1           5           10           15
Ala Val Ala Leu Pro Asn Gly Lys Ala Leu Leu Asp Met Ala Lys Arg
20           25           30
Ser Leu Thr Ser Ser Ala Glu Gly Glu Ser Asn Gly Tyr Phe Tyr Ser
35           40           45
Phe Trp Thr Asn Gly Glu Gly Asp Val Thr Tyr Thr Asn Gly Asp Ala
50           55           60
Gly Glu Tyr Thr Val Glu Trp Thr Asp Cys Gly Asp Phe Val Ala Gly
65           70           75           80
Lys Gly Trp Asn Pro Gly Ser Ala Gln Ala Val Thr Tyr Ser Gly Ser
85           90           95
Trp Glu Thr Asp Ala Asn Ala Tyr Leu Ser Val Tyr Gly Trp Thr Thr
100          105          110
Ser Pro Leu Val Glu Phe Tyr Ile Val Asp Lys Tyr Gly Asp Tyr Asp
115          120          125
Pro Ser Ser Gly Leu Thr Glu Leu Gly Thr Val Asp Ser Asp Asp Gly
130          135          140
Thr Tyr Lys Ile Tyr Gln Thr Thr Arg Glu Asp Ala Asp Ser Ile Glu
145          150          155          160
Gly Thr Ala Thr Phe Lys Gln Tyr Trp Ser Val Arg Thr Glu Gly Arg
165          170          175
Ile Gly Gly Thr Val Thr Thr Gln Asn His Phe Asp Ala Trp Glu Asn
180          185          190
Leu Gly Leu Glu Leu Gly Thr Phe Asn Tyr Met Ile Val Ala Thr Glu
195          200          205
Gly Tyr Glu Ser Ser Gly Ser Ala Thr Ile Thr Val Glu Thr Ser Ser
210          215          220
Asp Ser Thr Glu Glu Ser Cys
225          230

```

<210> 7

<211> 2460

<212> DNA

<213> Aspergillus niger

21062WO.ST25.txt

<400> 7

atacgtacca	gataagcagg	tgatttaatc	aatctcaaca	ccaaaggctg	ttcgaagcga	60
cggttaaagt	gatattcagg	cagtcaagat	tacaaactcc	tagcatgtat	atcaatagtt	120
tgtgagtttc	catgtggcgt	ttgaggcatt	tcctacgaaa	taggtggcct	gggggctaaa	180
tcaatattaa	tcacatgtag	ttatgttacg	cgcaggtgaa	agaaattagc	caggtagata	240
gacagcggga	ttcctccgac	cccttaatgt	cgcttaggtt	gtttcatagt	gctaaaattg	300
tagcctaaag	ggataaactg	cctgaggaat	atacaggttg	aaagatcggg	tctttgcaca	360
tccttccctt	tttcgtcgac	agacgaagca	tataaagggtg	agtgtctggc	aaacaaccgt	420
tgctatctca	acaccctcaa	ctttcaagtc	acagttgtac	ccttctactt	ggatatattc	480
tttcttatca	tcacatgggt	gtctttcctt	ggccaagcac	gcctggccgt	gccaatcctg	540
tcagcttttg	cgtgcatgct	tgcggaagt	tccgccatcc	ccccaccacc	gcgtggagct	600
ctttcacccg	aacgtctaca	atggattaga	gaggtgatcg	gcaatcagac	tgagaacgac	660
agcgtttcgg	accttgctaa	gcggagcact	atcctgcata	ccagccaaga	cggagttgac	720
agcgccggat	tctactactc	agtatacaat	gacaacgggg	ccgatgttg	atacaccgaa	780
tacccacga	ccggccagtt	cgaacttggc	tggagtgtctg	aggcggaatt	cctcgctgga	840
aagggcttca	agggcgga	cccacggtaa	gtagcccttc	accggccaga	gccaagaag	900
gctgacaatt	ttccggctca	gctccttgac	ctgggacggt	tatttcaccg	ccgaagggga	960
ctggactttg	gccatttatg	gctggaccac	aaaccccgtc	accgagtgg	atattgtgga	1020
gtcgcatgga	agtggaacc	caggcaacgg	gaacatcctc	ggccagggtg	atagtgatgg	1080
cggcgtctac	gatgtctata	atctccccta	tagaaatgtt	ccggagatct	acggtgtcac	1140
caacttcgac	cagcactgg	cggtgcgtcg	ctcccaccgc	tctactggca	ccgtcgacgt	1200
gagcgcccat	ttccagcgct	ggaaggaatt	gggtctcacc	cctggctccc	cagtcttcca	1260
gatggttacc	ctggaaggct	tttccggcca	gggttacctg	gacttcaccg	ttagtgcata	1320
gaaaggttca	gggacctcgt	ggtctcagct	ggcggcgagc	acttacttga	ggccagatcc	1380
cctccacgtg	tatctttttt	tctagggcaa	agaatggtga	cttcggttca	agaaacagaa	1440
gccccctttt	cccttccttt	ttgtttcttc	ttctcttaga	gcctgctctt	taacaacttg	1500
ttcgcttctt	ttccagttac	caattatttt	ttgaccttca	aagatattta	ctatcgaact	1560
aatttgttta	tcaacttcat	gcgttatgat	atatttcgaa	tctaaatgaa	gatcgtata	1620
ccaatgtgtc	atgtacaagg	taatctcggc	agcacagatg	tttttgatta	ttgatgttgt	1680
gatatcttac	tactttaacc	ctggaatgct	cttctccggg	aagaatctcg	taatggacct	1740
agaacgcgg	agaccagggg	tctagggacc	tagagccgc	ttgataaaat	gtagtcttgc	1800
gatccaataa	gcaatctcct	aaagcttact	tatctgctaa	attcataatg	agttacagct	1860
agtatttaga	aatatatatt	ttatagttat	aaaaacagta	atatagcaag	cttcataag	1920
aaaggatcat	agttgctctg	gattgaattt	gccagtaaat	aaaatggcat	aacgaccgcg	1980
gtacacttaa	tgggttttaa	taaaagattc	aaagtttgac	ttaaattcat	atactgatcg	2040
ctgcccctct	atcttgact	gttgtttaac	tgtttaggcg	ggcttgacgg	taaaactatg	2100
atccgactca	atacatttca	ctaatgaata	accattggct	ggatccccgc	ttcgcgctcg	2160
aaagatatct	ggacattcac	actcgaaatt	ctgtctcgtc	tgaacggacc	acggcttttt	2220
aatgcgccgg	acgtgtcttt	ttctgcgcac	aagtctaaat	cactataatc	ttggttcttg	2280
atacccaagc	atcgtattgg	tgttctctct	gtattacttt	aggcgccacc	agaaccctcc	2340
acagccatct	agaaagactc	ttgaggagat	gtcccgatct	cttgggatgt	tgtacgctgg	2400
ctctcaacag	cttttcaaaa	ttacatactg	attgataatc	ttagtgccct	gagtgccgga	2460

21062W0.ST25.txt

<210> 8
 <211> 771
 <212> DNA
 <213> *Aspergillus niger*

<220>
 <221> CDS
 <222> (1)..(771)

<400> 8
 atg gtg tct ttc ctt ggc caa gca cgc ctg gcc gtg cca atc ctg tca 48
 Met Val Ser Phe Leu Gly Gln Ala Arg Leu Ala Val Pro Ile Leu Ser
 1 5 10 15
 gct ttt gcg tgc atg ctt gcg gca agt tcc gcc atc ccc cca cca ccg 96
 Ala Phe Ala Cys Met Leu Ala Ala Ser Ser Ala Ile Pro Pro Pro Pro
 20 25 30
 cgt gga gct ctt tca ccc gaa cgt cta caa tgg att aga gag gtg atc 144
 Arg Gly Ala Leu Ser Pro Glu Arg Leu Gln Trp Ile Arg Glu Val Ile
 35 40 45
 ggc aat cag act gag aac gac agc gtt tcg gac ctt gct aag cgg agc 192
 Gly Asn Gln Thr Glu Asn Asp Ser Val Ser Asp Leu Ala Lys Arg Ser
 50 55 60
 act atc ctg cat acc agc caa gac gga gtt gac agc gcc gga ttc tac 240
 Thr Ile Leu His Thr Ser Gln Asp Gly Val Asp Ser Ala Gly Phe Tyr
 65 70 75 80
 tac tca gta tac aat gac aac ggg gcc gat gtt gga tac acc gaa tac 288
 Tyr Ser Val Tyr Asn Asp Asn Gly Ala Asp Val Gly Tyr Thr Glu Tyr
 85 90 95
 ccc acg acc ggc cag ttc gaa ctt ggc tgg agt gct gag gcg gaa ttc 336
 Pro Thr Thr Gly Gln Phe Glu Leu Gly Trp Ser Ala Glu Ala Glu Phe
 100 105 110
 ctc gct gga aag ggc ttc aag ggc ggc aac cca cgc tcc ttg acc tgg 384
 Leu Ala Gly Lys Gly Phe Lys Gly Gly Asn Pro Arg Ser Leu Thr Trp
 115 120 125
 gac ggt tat ttc acc gcc gaa ggg gac tgg act ttg gcc att tat ggc 432
 Asp Gly Tyr Phe Thr Ala Glu Gly Asp Trp Thr Leu Ala Ile Tyr Gly
 130 135 140
 tgg acc aca aac ccc gtc acc gag tgg tat att gtg gag tcg cat gga 480
 Trp Thr Thr Asn Pro Val Thr Glu Trp Tyr Ile Val Glu Ser His Gly
 145 150 155 160
 agt gga acc cca ggc aac ggg aac atc ctc ggc cag gtt gat agt gat 528
 Ser Gly Thr Pro Gly Asn Gly Asn Ile Leu Gly Gln Val Asp Ser Asp
 165 170 175

21062W0.ST25.txt

```

ggc ggc gtc tac gat gtc tat aat ctc ccc tat aga aat gtt ccg gag      576
Gly Gly Val Tyr Asp Val Tyr Asn Leu Pro Tyr Arg Asn Val Pro Glu
      180                      185                      190
atc tac ggt gtc acc aac ttc gac cag cac tgg tcg gtg cgt cgc tcc      624
Ile Tyr Gly Val Thr Asn Phe Asp Gln His Trp Ser Val Arg Arg Ser
      195                      200                      205
cac cgc tct act ggc acc gtc gac gtg agc gcc cat ttc cag cgc tgg      672
His Arg Ser Thr Gly Thr Val Asp Val Ser Ala His Phe Gln Arg Trp
      210                      215                      220
aag gaa ttg ggt ctc acc cct ggc tcc cca gtc ttc cag atg gtt acc      720
Lys Glu Leu Gly Leu Thr Pro Gly Ser Pro Val Phe Gln Met Val Thr
      225                      230                      235                      240
ctg gaa ggc ttt tcc ggc cag ggg tac ctg gac ttc acc gtt agt gca      768
Leu Glu Gly Phe Ser Gly Gln Gly Tyr Leu Asp Phe Thr Val Ser Ala
      245                      250                      255
tag                                                                    771

```

<210> 9

<211> 256

<212> PRT

<213> Aspergillus niger

<400> 9

```

Met Val Ser Phe Leu Gly Gln Ala Arg Leu Ala Val Pro Ile Leu Ser
1          5          10          15
Ala Phe Ala Cys Met Leu Ala Ala Ser Ser Ala Ile Pro Pro Pro Pro
      20          25          30
Arg Gly Ala Leu Ser Pro Glu Arg Leu Gln Trp Ile Arg Glu Val Ile
      35          40          45
Gly Asn Gln Thr Glu Asn Asp Ser Val Ser Asp Leu Ala Lys Arg Ser
      50          55          60
Thr Ile Leu His Thr Ser Gln Asp Gly Val Asp Ser Ala Gly Phe Tyr
65          70          75          80
Tyr Ser Val Tyr Asn Asp Asn Gly Ala Asp Val Gly Tyr Thr Glu Tyr
      85          90          95
Pro Thr Thr Gly Gln Phe Glu Leu Gly Trp Ser Ala Glu Ala Glu Phe
      100         105         110
Leu Ala Gly Lys Gly Phe Lys Gly Gly Asn Pro Arg Ser Leu Thr Trp
      115         120         125
Asp Gly Tyr Phe Thr Ala Glu Gly Asp Trp Thr Leu Ala Ile Tyr Gly
      130         135         140
Trp Thr Thr Asn Pro Val Thr Glu Trp Tyr Ile Val Glu Ser His Gly
145         150         155         160

```

21062W0.ST25.txt

Ser Gly Thr Pro Gly Asn Gly Asn Ile Leu Gly Gln Val Asp Ser Asp
 165 170 175
 Gly Gly Val Tyr Asp Val Tyr Asn Leu Pro Tyr Arg Asn Val Pro Glu
 180 185 190
 Ile Tyr Gly Val Thr Asn Phe Asp Gln His Trp Ser Val Arg Arg Ser
 195 200 205
 His Arg Ser Thr Gly Thr Val Asp Val Ser Ala His Phe Gln Arg Trp
 210 215 220
 Lys Glu Leu Gly Leu Thr Pro Gly Ser Pro Val Phe Gln Met Val Thr
 225 230 235 240
 Leu Glu Gly Phe Ser Gly Gln Gly Tyr Leu Asp Phe Thr Val Ser Ala
 245 250 255

<210> 10

<211> 3306

<212> DNA

<213> *Aspergillus niger*

<400> 10

```

aggaggcggc atattacacc ataggctttt cagacaaaat tactgcagaa gtctcagcca      60
cgcaatgtta ccatgccacg ccaatcaggt cagctgcttg aattgggttc agaatcccca      120
ggaacccgag cggcccagag ctgccgagag tcacagacga tgcggagggc cacttggcat      180
ccaatgtctg cgttttatga tttccagagc atgtctatct tgaccatata agccacaccg      240
tagtgctatg tcatagactg tctatcgga gatgtccagt aattgtccgt gtttcttaca      300
gatgctcggg aagtcaatag actgcagggc tcttcatgcg tcatgatata accgacggca      360
cctcctttgg gctgtgctta gtgggtccgg taagctagac attcaagcgc agccattcac      420
agcttcagtc caaagacatg accgaatagt agtatcctaa tagagttcga gtagcatgac      480
cacatggtgt gaagaacatg gggtccctc acggcgccgg ttaaaagcca tccctcctc      540
aaagaagaaa catcgtacga cgtacacggg tcggttggtt caactattat atttcattct      600
aactatactt attgcaaagc ggaatttggt ccgataatca tacaaggctt tcaggggtag      660
tcccatctca atatccgggg taatgacatc caagtcgtac ctaccaaacc ttacctatta      720
ctcgaaagta gcaactgctc cggaaaagca ttcaccaacc ctccacaaaa ctaacttcca      780
caggaatgcg gcacctccat tccctccaca attaccgagg cttatcccag gaggaacaga      840
tccctgaaga aacgtgacga accgtgatga gcgggcaatc acagattagc ctcttcccc      900
gcatcccccg acgggcaaaa ccccggtggc actggcatta gccctaggca cggacaaagc      960
cattgtcaag ccttcaggaa ctaaataaac taaatctgcc catccctgcg gggtagatac      1020
tccacgacgc cctgcacgcc cgccggaatg ccgtcgtcgc ttgacatcat gtggaacaac      1080
acattccttc tgcctcaaaa ctgagccttt taaccgaccg gttacttaac aatgtataca      1140
ggagggtctca atattggagg aaaatccctg acccctgaag ccacacatgt aactagactt      1200
ttgggatctt gcatttcatg cataggcact gactgcagta aagaatataa ataagcctcc      1260
cttcgcaccc tagtagcatg atttctttct ccaagcggcc tatctcactg tctcctctcc      1320
tagcccagaa cttactgagc aggcagtaat cctccacat gttctccgc cgaaacctcg      1380
tcgccctagg gctggcagcc accgtcagcg ccggcccctg tgacatctac gaagccggcg      1440

```

21062W0.ST25.txt

acacgccctg	cgtagccgcg	cacagcacca	cccgcgccct	atacagctcc	ttcagcggcg	1500
ccctctatca	gctccaacgt	ggctccgacg	acaccaccac	caccatctcc	ccgctcacag	1560
ccggtggcgt	cgccgacgcc	tccgcccgaag	acaccttctg	tgccaacacc	acctgcctga	1620
tcaccatcat	ctacgaccaa	tccggcaacg	gcaaccacct	gaccaagca	cccccgagg	1680
gcttcgacgg	ccccgacgtc	gacggctacg	acaacctcgc	cagcgccatc	ggcgcccccg	1740
tcaccctcaa	tggccaaaag	gcctacggcg	tgttcattgtc	ccccggcacg	ggctaccgta	1800
acaacgaagc	caccggcact	gccaccggcg	acgagcccga	aggcatgtac	gccgtcctgg	1860
acggcacgca	ctacaacgac	gcctgctgct	tcgactacgg	caacgccgag	accagcagca	1920
ccgacactgg	cgccggccac	atggaggcca	tctacctcgg	gaacagcacc	acctgggggt	1980
acggcgccgg	cgacggggcc	tggatcatgg	tcgatattga	gaacaacctc	ttctccgggtg	2040
ctgatgaggg	atataactcc	ggtgatccct	cgatcagcta	cagctttgtc	actgccgcgg	2100
tcaagggcgg	ggctgataag	tgggctattc	gcggtggtaa	tgctgcctct	gggtccctct	2160
ctacttacta	cagcggcgct	cgcccggtat	actccggcta	taaccccatg	agcaaggagg	2220
gcgctatcat	cctgggtatc	ggcggtgaca	acagcaacgg	cgcccagggt	accttctacg	2280
aggggtgtcat	gacctccggc	taccgctcgg	acgatgtcga	gaactccgtc	caggagaaca	2340
tcgtggctgc	gaaatacgtc	tccggctcgc	tggtcagcgg	cccgtcgttc	acctccggag	2400
aggtggtctc	gctgcgtgtc	actacgcccg	gttacacgac	gcggtatatt	gcgcacactg	2460
acaccactgt	gaacacgcag	gtcgtggacg	acgatagtct	caccacgctg	aaggaggagg	2520
ctagctggac	cgtggtgacc	ggtctggcta	atagtcagtg	cttctcgttc	gagtcggttg	2580
atacccctgg	tagctataat	cggcattaca	actttgagtt	gctgcttaat	gccaatgatg	2640
gcacgaagca	gttccatgag	gatgtactct	tctgtcctca	ggcgccgttg	aatggagagg	2700
gtacttcgtt	gagatcgtgg	agttatccga	cgaggatatt	cagacattat	gataatgtcc	2760
tgtatgctgc	tagtaatggc	ggtgtgcaga	cgtttgattc	caagacatcg	tttaataatg	2820
atgttagctt	tgagattgag	acggcgtttg	cttcgtaagg	ggggaattgg	gggttggttg	2880
gtggtttggg	tgaggggttg	tatctgcctg	tgttggcggg	gagagtgggt	tgtaaataag	2940
ttatTTTTTg	tatatagcac	atgaatacta	tacgatagta	tacatgatag	tggttggttc	3000
atgtaatcga	gaaattattc	aaggcttgca	atacaatata	ttattactgt	atctcgtgga	3060
gcattcacag	actgaacggc	tgcaaatga	ttcttatgcg	gtgatattga	ctgtgattag	3120
atgatatatt	caggtgtgct	atTTTTcttg	tatttatgct	tcatgcgaat	atatgtattt	3180
cacgacaggc	agagagggtta	aacaaagcat	actaacgagc	gattcagata	tgacatgtat	3240
gaataagaca	aaagtactgc	gacactcaag	accaacacaa	accagtatac	ttcataacaa	3300
tatatg						3306

<210> 11

<211> 1500

<212> DNA

<213> *Aspergillus niger*

<220>

<221> CDS

<222> (1)..(1500)

<400> 11

21062W0.ST25.txt

atg ttc tcc cgc cga aac ctc gtc gcc cta ggg ctg gca gcc acc gtc	48
Met Phe Ser Arg Arg Asn Leu Val Ala Leu Gly Leu Ala Ala Thr Val	
1 5 10 15	
agc gcc ggc ccc tgt gac atc tac gaa gcc ggc gac acg ccc tgc gta	96
Ser Ala Gly Pro Cys Asp Ile Tyr Glu Ala Gly Asp Thr Pro Cys Val	
20 25 30	
gcc gcg cac agc acc acc cgc gcc cta tac agc tcc ttc agc ggc gcc	144
Ala Ala His Ser Thr Thr Arg Ala Leu Tyr Ser Ser Phe Ser Gly Ala	
35 40 45	
ctc tat cag ctc caa cgt ggc tcc gac gac acc acc acc acc atc tcc	192
Leu Tyr Gln Leu Gln Arg Gly Ser Asp Asp Thr Thr Thr Thr Ile Ser	
50 55 60	
ccg ctc aca gcc ggt ggc gtc gcc gac gcc tcc gcc caa gac acc ttc	240
Pro Leu Thr Ala Gly Gly Val Ala Asp Ala Ser Ala Gln Asp Thr Phe	
65 70 75 80	
tgt gcc aac acc acc tgc ctg atc acc atc atc tac gac caa tcc ggc	288
Cys Ala Asn Thr Thr Cys Leu Ile Thr Ile Ile Tyr Asp Gln Ser Gly	
85 90 95	
aac ggc aac cac ctg acc caa gca ccc ccg gga ggc ttc gac ggc ccc	336
Asn Gly Asn His Leu Thr Gln Ala Pro Pro Gly Gly Phe Asp Gly Pro	
100 105 110	
gac gtc gac ggc tac gac aac ctc gcc agc gcc atc ggc gcc ccc gtc	384
Asp Val Asp Gly Tyr Asp Asn Leu Ala Ser Ala Ile Gly Ala Pro Val	
115 120 125	
acc ctc aat ggc caa aag gcc tac ggc gtg ttc atg tcc ccc ggc acg	432
Thr Leu Asn Gly Gln Lys Ala Tyr Gly Val Phe Met Ser Pro Gly Thr	
130 135 140	
ggc tac cgt aac aac gaa gcc acc ggc act gcc acc ggc gac gag ccc	480
Gly Tyr Arg Asn Asn Glu Ala Thr Gly Thr Ala Thr Gly Asp Glu Pro	
145 150 155 160	
gaa ggc atg tac gcc gtc ctg gac ggc acg cac tac aac gac gcc tgc	528
Glu Gly Met Tyr Ala Val Leu Asp Gly Thr His Tyr Asn Asp Ala Cys	
165 170 175	
tgc ttc gac tac ggc aac gcc gag acc agc agc acc gac act ggc gcc	576
Cys Phe Asp Tyr Gly Asn Ala Glu Thr Ser Ser Thr Asp Thr Gly Ala	
180 185 190	
ggc cac atg gag gcc atc tac ctc ggg aac agc acc acc tgg ggc tac	624
Gly His Met Glu Ala Ile Tyr Leu Gly Asn Ser Thr Thr Trp Gly Tyr	
195 200 205	
ggc gcc ggc gac ggg ccc tgg atc atg gtc gat atg gag aac aac ctc	672
Gly Ala Gly Asp Gly Pro Trp Ile Met Val Asp Met Glu Asn Asn Leu	
210 215 220	
ttc tcc ggt gct gat gag gga tat aac tcc ggt gat ccc tcg atc agc	720
Phe Ser Gly Ala Asp Glu Gly Tyr Asn Ser Gly Asp Pro Ser Ile Ser	

21062W0.ST25.txt

225	230	235	240	
tac agc ttt gtc act gcc gcg gtc aag ggc ggg gct gat aag tgg gct				768
Tyr Ser Phe Val Thr Ala Ala Val Lys Gly Gly Ala Asp Lys Trp Ala				
	245	250	255	
att cgc ggt ggt aat gct gcc tct ggg tcc ctc tct act tac tac agc				816
Ile Arg Gly Gly Asn Ala Ala Ser Gly Ser Leu Ser Thr Tyr Tyr Ser				
	260	265	270	
ggc gct cgc ccg gat tac tcc ggc tat aac ccc atg agc aag gag ggc				864
Gly Ala Arg Pro Asp Tyr Ser Gly Tyr Asn Pro Met Ser Lys Glu Gly				
	275	280	285	
gct atc atc ctg ggt atc ggc ggt gac aac agc aac ggc gcc cag ggt				912
Ala Ile Ile Leu Gly Ile Gly Gly Asp Asn Ser Asn Gly Ala Gln Gly				
	290	295	300	
acc ttc tac gag ggt gtc atg acc tcc ggc tac ccg tcg gac gat gtc				960
Thr Phe Tyr Glu Gly Val Met Thr Ser Gly Tyr Pro Ser Asp Asp Val				
	305	310	315	320
gag aac tcc gtc cag gag aac atc gtg gct gcg aaa tac gtc tcc ggc				1008
Glu Asn Ser Val Gln Glu Asn Ile Val Ala Ala Lys Tyr Val Ser Gly				
	325	330	335	
tcg ctg gtc agc ggc ccg tcg ttc acc tcc gga gag gtg gtc tcg ctg				1056
Ser Leu Val Ser Gly Pro Ser Phe Thr Ser Gly Glu Val Val Ser Leu				
	340	345	350	
cgt gtc act acg ccc ggt tac acg acg cgg tat att gcg cac act gac				1104
Arg Val Thr Thr Pro Gly Tyr Thr Thr Arg Tyr Ile Ala His Thr Asp				
	355	360	365	
acc act gtg aac acg cag gtc gtg gac gac gat agt tcc acc acg ctg				1152
Thr Thr Val Asn Thr Gln Val Val Asp Asp Asp Ser Ser Thr Thr Leu				
	370	375	380	
aag gag gag gct agc tgg acc gtg gtg acc ggt ctg gct aat agt cag				1200
Lys Glu Glu Ala Ser Trp Thr Val Val Thr Gly Leu Ala Asn Ser Gln				
	385	390	395	400
tgc ttc tcg ttc gag tcg gtt gat acc cct ggt agc tat atc cgg cat				1248
Cys Phe Ser Phe Glu Ser Val Asp Thr Pro Gly Ser Tyr Ile Arg His				
	405	410	415	
tac aac ttt gag ttg ctg ctt aat gcc aat gat ggc acg aag cag ttc				1296
Tyr Asn Phe Glu Leu Leu Leu Asn Ala Asn Asp Gly Thr Lys Gln Phe				
	420	425	430	
cat gag gat gct act ttc tgt cct cag gcg ccg ttg aat gga gag ggt				1344
His Glu Asp Ala Thr Phe Cys Pro Gln Ala Pro Leu Asn Gly Glu Gly				
	435	440	445	
act tcg ttg aga tcg tgg agt tat ccg acg agg tat ttc aga cat tat				1392
Thr Ser Leu Arg Ser Trp Ser Tyr Pro Thr Arg Tyr Phe Arg His Tyr				
	450	455	460	
gat aat gtc ctg tat gct gct agt aat ggc ggt gtg cag acg ttt gat				1440

21062W0.ST25.txt

Asp Asn Val Leu Tyr Ala Ala Ser Asn Gly Gly Val Gln Thr Phe Asp
 465 470 475 480
 tcc aag aca tcg ttt aat aat gat gtt agc ttt gag att gag acg gcg 1488
 Ser Lys Thr Ser Phe Asn Asn Asp Val Ser Phe Glu Ile Glu Thr Ala
 485 490 495
 ttt gct tcg taa 1500
 Phe Ala Ser

<210> 12

<211> 499

<212> PRT

<213> Aspergillus niger

<400> 12

Met Phe Ser Arg Arg Asn Leu Val Ala Leu Gly Leu Ala Ala Thr Val
 1 5 10 15
 Ser Ala Gly Pro Cys Asp Ile Tyr Glu Ala Gly Asp Thr Pro Cys Val
 20 25 30
 Ala Ala His Ser Thr Thr Arg Ala Leu Tyr Ser Ser Phe Ser Gly Ala
 35 40 45
 Leu Tyr Gln Leu Gln Arg Gly Ser Asp Asp Thr Thr Thr Thr Ile Ser
 50 55 60
 Pro Leu Thr Ala Gly Gly Val Ala Asp Ala Ser Ala Gln Asp Thr Phe
 65 70 75 80
 Cys Ala Asn Thr Thr Cys Leu Ile Thr Ile Ile Tyr Asp Gln Ser Gly
 85 90 95
 Asn Gly Asn His Leu Thr Gln Ala Pro Pro Gly Gly Phe Asp Gly Pro
 100 105 110
 Asp Val Asp Gly Tyr Asp Asn Leu Ala Ser Ala Ile Gly Ala Pro Val
 115 120 125
 Thr Leu Asn Gly Gln Lys Ala Tyr Gly Val Phe Met Ser Pro Gly Thr
 130 135 140
 Gly Tyr Arg Asn Asn Glu Ala Thr Gly Thr Ala Thr Gly Asp Glu Pro
 145 150 155 160
 Glu Gly Met Tyr Ala Val Leu Asp Gly Thr His Tyr Asn Asp Ala Cys
 165 170 175
 Cys Phe Asp Tyr Gly Asn Ala Glu Thr Ser Ser Thr Asp Thr Gly Ala
 180 185 190
 Gly His Met Glu Ala Ile Tyr Leu Gly Asn Ser Thr Thr Trp Gly Tyr
 195 200 205
 Gly Ala Gly Asp Gly Pro Trp Ile Met Val Asp Met Glu Asn Asn Leu
 210 215 220
 Phe Ser Gly Ala Asp Glu Gly Tyr Asn Ser Gly Asp Pro Ser Ile Ser

225	230							235							240		
Tyr	Ser	Phe	Val	Thr	Ala	Ala	Val	Lys	Gly	Gly	Ala	Asp	Lys	Trp	Ala		
				245				250				255					
Ile	Arg	Gly	Gly	Asn	Ala	Ala	Ser	Gly	Ser	Leu	Ser	Thr	Tyr	Tyr	Ser		
				260				265				270					
Gly	Ala	Arg	Pro	Asp	Tyr	Ser	Gly	Tyr	Asn	Pro	Met	Ser	Lys	Glu	Gly		
				275				280				285					
Ala	Ile	Ile	Leu	Gly	Ile	Gly	Gly	Asp	Asn	Ser	Asn	Gly	Ala	Gln	Gly		
				290				295				300					
Thr	Phe	Tyr	Glu	Gly	Val	Met	Thr	Ser	Gly	Tyr	Pro	Ser	Asp	Asp	Val		
305				310				315				320					
Glu	Asn	Ser	Val	Gln	Glu	Asn	Ile	Val	Ala	Ala	Lys	Tyr	Val	Ser	Gly		
				325				330				335					
Ser	Leu	Val	Ser	Gly	Pro	Ser	Phe	Thr	Ser	Gly	Glu	Val	Val	Ser	Leu		
				340				345				350					
Arg	Val	Thr	Thr	Pro	Gly	Tyr	Thr	Thr	Arg	Tyr	Ile	Ala	His	Thr	Asp		
				355				360				365					
Thr	Thr	Val	Asn	Thr	Gln	Val	Val	Asp	Asp	Asp	Ser	Ser	Thr	Thr	Leu		
				370				375				380					
Lys	Glu	Glu	Ala	Ser	Trp	Thr	Val	Val	Thr	Gly	Leu	Ala	Asn	Ser	Gln		
385				390				395				400					
Cys	Phe	Ser	Phe	Glu	Ser	Val	Asp	Thr	Pro	Gly	Ser	Tyr	Ile	Arg	His		
				405				410				415					
Tyr	Asn	Phe	Glu	Leu	Leu	Leu	Asn	Ala	Asn	Asp	Gly	Thr	Lys	Gln	Phe		
				420				425				430					
His	Glu	Asp	Ala	Thr	Phe	Cys	Pro	Gln	Ala	Pro	Leu	Asn	Gly	Glu	Gly		
				435				440				445					
Thr	Ser	Leu	Arg	Ser	Trp	Ser	Tyr	Pro	Thr	Arg	Tyr	Phe	Arg	His	Tyr		
450				455				460									
Asp	Asn	Val	Leu	Tyr	Ala	Ala	Ser	Asn	Gly	Gly	Val	Gln	Thr	Phe	Asp		
465				470				475				480					
Ser	Lys	Thr	Ser	Phe	Asn	Asn	Asp	Val	Ser	Phe	Glu	Ile	Glu	Thr	Ala		
				485				490				495					
Phe Ala Ser																	

<213> *Aspergillus niger*

tgtggaaggg agtagatatg catacattat tggaaagaca ataagcttaa gctacaggcc
tagaccccca gctaagatct cgggatggat cctctccacg atcaccacaac ccagcctcct

21062W0.ST25.txt

aaatactctt	atcatcgacc	cgacattcct	gtaaagcaac	taaagctccc	cacccaactg	180
gatgtctcat	ccagtatctc	aaccctagta	accctatcct	cgccctgaat	tgacaccacc	240
tcatcatgaa	cctttcatac	aggcctcccg	tcaacagata	accccatgat	cctggatctt	300
cgccaccaac	tccaccttag	atgtcaaaat	tggggcctct	aatatcggcg	agcgtctgtg	360
tgtttaggac	ttatgataag	ccaagtcggg	gagggttaagg	cggtaggtta	tatcgagggt	420
ttggtgggtg	gggagcagag	ggtgaggcgg	gaaggttcta	gacaatcgac	atttgttggc	480
taggtggtat	gtttcacaga	tattgattct	aatcaaactg	cccagtcgta	ataggagatt	540
gtggataatc	gatagctaca	ggcaacagaa	atctgcggtc	aagaggatac	tgtaggcagg	600
agataactct	gggtcagaga	ggactggggg	cagaagaaaa	gtgcaggcac	gagagaacaa	660
ctgcagctac	tatgagacta	cagagaagaa	attaccgcag	ccttagttgg	attgcggccc	720
ccactgacat	ggccccatta	cccctgggga	ccagagaaca	gtaatatgat	aaacacatga	780
tgtactagag	tcattccgtg	ggacccgagc	aatccgcccc	acccgagcac	cccaaaacaa	840
cgggttctga	ccaacacgtc	cctgcctgac	caccaccacc	accactaccg	ccaccaccaa	900
tcttgccaat	tattgtctaa	atcagcgggt	tttgactcta	ttcaatgtct	agaatccggc	960
taagatcaaa	tatctatatc	tagagacgtt	gaaaaatgag	aaacgaaaat	gactcctatc	1020
gtacttcctc	gcctgtgcbt	tgaattactc	cgccacacac	tgtgtgaatc	ccgcaagatg	1080
cacaagtctc	gcbtgttca	acgccaggag	tttaagaagg	gcatgtttat	acgagcagat	1140
agcaagcttc	ttaccgctga	gcttcgtatc	agtcaaaaca	cctaagccct	gttgatatatt	1200
caccgacaaa	aggccgatcc	ctaagccatg	ctcatctaaa	ttgagttggt	ctatccttct	1260
aaattaatat	cgtcagcaca	agcaggaccc	ttttcttggt	tcccatcaga	tcccgttctc	1320
cacattgttc	ggttaccgac	tgccacggac	catcactaac	tttaaacgga	aagtgtctccg	1380
atatagggtc	tcattttgaa	acggaggatg	tgcatgtctg	ttgcagattg	gcgtgagatg	1440
gcggcgcaac	ccgcaatagc	tagttggcct	acttcttctt	gaatgccata	ttggacacgg	1500
aaattcggac	aatcctaatt	gcgatagtaa	gcattgttga	gagattcagg	tatttaagct	1560
gcctcatctg	ccgtccctc	cgtgttggtg	tcttatccta	aacactgtc	ctcttattcc	1620
tttcgaattc	cccttgattg	ccgtcctcc	gctctaacgc	ccaacatgaa	gctccctgtg	1680
acatttgcta	tgcttgccgc	caccgccatg	ggccagacga	tgtgctctca	atatgacagt	1740
gcctcgagcc	ccccatatc	agtgaaccag	aacctctggg	gcgagtacca	aggcaccggc	1800
agccagtgtg	tatatgtcga	caaactctcc	agcagtgggt	catcctggca	caccgaatgg	1860
acctggagcg	gtggtgaggg	aacagtgaat	agctactcta	actctggcgt	tacatttaac	1920
aagaagctcg	tgagtgatgt	atcaagcatc	cccacctcgg	tggaaatgga	gcaggacaac	1980
accaacgtca	acgccgatgt	cgcgtatgat	cttttcaccg	cggcgaatgt	ggaccatgct	2040
acttctagcg	gcgactatga	actgatgatt	tggtatgtga	gctcgcaaag	aatcaaagaa	2100
tacagaaatc	taacaacccc	aggcttgccc	gctacggcaa	catccagccc	attggcaagc	2160
aaattgccac	ggccacagtg	ggaggcaagt	cctgggagggt	gtggtatggc	agcaccaccc	2220
aggccggtgc	ggagcagagg	acatacagct	tcgtgtcggga	aagccctatc	aactcataca	2280
gtggggacat	caatgcattt	ttcagctatc	taactcagaa	ccaaggcttt	cccgccagct	2340
ctcagtactt	gatcagtaag	attccctgat	tctaccagcg	gacgccgcag	cgaaacagtc	2400
actaacagaa	gtgatgtgta	gatctgcagt	ttggaactga	ggcgttcacc	gggggcccgg	2460
caaccttcac	ggttgacaac	tggaccgcca	gtgtcaaacta	gggttctaga	agtagccttt	2520
gaggcagaat	ctgggtaaat	tgactccagc	tcgggataat	gattgcttgt	ttcttcgttc	2580
tggaaagtgt	ggcgtgtgag	agctaaaaag	tcgtacccac	tctgattgga	aagacttatt	2640
caacattggt	ccttcccttc	tgttgggcaa	ggcatagtta	gtgattagac	aagtcaaggt	2700
catggtggat	cccctgtatt	cctagccagc	cttttcctag	aatggtaatc	cacccgtagg	2760

21062W0.ST25.txt

```

gggatgatca ttaacaaatg ccagtcgtac taggttaa at tgtaaccga tcccgtaca 2820
gtagctagcg catccggaag acccattctt agtgggctag gtatttgta agcactgtag 2880
aaaggaacag atcaactcct gggaatttca aaatagcata agcagatacc tgccccgagc 2940
attttacctt gggatcactt tcctcttcac ggtcccaggc ggtggactct ctaaggaaga 3000
ggaagcattg cttgccaatg tccggtagga atgtcataac gtccggattg gcgactcgat 3060
gttccatgct atacaaggcc tgcaacactg ggttgaatcc acggcttata gttttccacg 3120
gcttgtcatt ttccgcggt tgtgggttta gcttctacac ttcgtcacgg ctccgaatcg 3180
gaccatgctg ggatcttcgg catataatcc agtatatcaa agccgatagg gacaggggtt 3240
aaattagagg gatgtcacta gacagggatc tgtaagacgt ccggccttga ttgtaggctc 3300
caattttaac tattcccggg aacggcaatt ggttgccctg aatgtcaagg cgttttgacc 3360
cgtactctat ccgcagcggg ggctgcaata tggcttttac cgatggaatt aagacttatt 3420
attttccga cgtactagtt gctccactat gagctgttct aaaaagtcga aggtaggcag 3480
tcatcaaaca gaatgatgtt ataatagcct ataactgccg agattaccct gagtttcacc 3540
tgtccagata tggcccaata atatcaggct aaatgaatct tcccctgttg gcactttcac 3600
cgcacaccct aatttgggg 3619

```

<210> 14

<211> 720

<212> DNA

<213> *Aspergillus niger*

<220>

<221> CDS

<222> (1)..(720)

<400> 14

```

atg aag ctc cct gtg aca ctt gct atg ctt gcg gcc acc gcc atg ggc 48
Met Lys Leu Pro Val Thr Leu Ala Met Leu Ala Ala Thr Ala Met Gly
1 5 10 15
cag acg atg tgc tct caa tat gac agt gcc tcg agc ccc cca tat tca 96
Gln Thr Met Cys Ser Gln Tyr Asp Ser Ala Ser Ser Pro Pro Tyr Ser
20 25 30
gtg aac cag aac ctc tgg ggc gag tac caa ggc acc ggc agc cag tgt 144
Val Asn Gln Asn Leu Trp Gly Glu Tyr Gln Gly Thr Gly Ser Gln Cys
35 40 45
gta tat gtc gac aaa ctc tcc agc agt ggt gca tcc tgg cac acc gaa 192
Val Tyr Val Asp Lys Leu Ser Ser Ser Gly Ala Ser Trp His Thr Glu
50 55 60
tgg acc tgg agc ggt ggt gag gga aca gtg aaa agc tac tct aac tct 240
Trp Thr Trp Ser Gly Gly Glu Gly Thr Val Lys Ser Tyr Ser Asn Ser
65 70 75 80
ggc gtt aca ttt aac aag aag ctc gtg agt gat gta tca agc atc ccc 288
Gly Val Thr Phe Asn Lys Lys Leu Val Ser Asp Val Ser Ser Ile Pro

```

21062W0.ST25.txt

85	90	95	
acc tcg gtg gaa tgg aag cag gac aac acc aac gtc aac gcc gat gtc			336
Thr Ser Val Glu Trp Lys Gln Asp Asn Thr Asn Val Asn Ala Asp Val			
100	105	110	
gcg tat gat ctt ttc acc gcg gcg aat gtg gac cat gct act tct agc			384
Ala Tyr Asp Leu Phe Thr Ala Ala Asn Val Asp His Ala Thr Ser Ser			
115	120	125	
ggc gac tat gaa ctg atg att tgg ctt gcc cgc tac ggc aac atc cag			432
Gly Asp Tyr Glu Leu Met Ile Trp Leu Ala Arg Tyr Gly Asn Ile Gln			
130	135	140	
ccc att ggc aag caa att gcc acg gcc aca gtg gga ggc aag tcc tgg			480
Pro Ile Gly Lys Gln Ile Ala Thr Ala Thr Val Gly Gly Lys Ser Trp			
145	150	155	160
gag gtg tgg tat ggc agc acc acc cag gcc ggt gcg gag cag agg aca			528
Glu Val Trp Tyr Gly Ser Thr Thr Gln Ala Gly Ala Glu Gln Arg Thr			
165	170	175	
tac agc ttc gtg tcg gaa agc cct atc aac tca tac agt ggg gac atc			576
Tyr Ser Phe Val Ser Glu Ser Pro Ile Asn Ser Tyr Ser Gly Asp Ile			
180	185	190	
aat gca ttt ttc agc tat cta act cag aac caa ggc ttt ccc gcc agc			624
Asn Ala Phe Phe Ser Tyr Leu Thr Gln Asn Gln Gly Phe Pro Ala Ser			
195	200	205	
tct cag tac ttg atc aat ctg cag ttt gga act gag gcg ttc acc ggg			672
Ser Gln Tyr Leu Ile Asn Leu Gln Phe Gly Thr Glu Ala Phe Thr Gly			
210	215	220	
ggc ccg gca acc ttc acg gtt gac aac tgg acc gcc agt gtc aac tag			720
Gly Pro Ala Thr Phe Thr Val Asp Asn Trp Thr Ala Ser Val Asn			
225	230	235	

<210> 15

<211> 239

<212> PRT

<213> Aspergillus niger

<400> 15

Met Lys Leu Pro Val Thr Leu Ala Met Leu Ala Ala Thr Ala Met Gly	
1	15
Gln Thr Met Cys Ser Gln Tyr Asp Ser Ala Ser Ser Pro Pro Tyr Ser	
20	30
Val Asn Gln Asn Leu Trp Gly Glu Tyr Gln Gly Thr Gly Ser Gln Cys	
35	45
Val Tyr Val Asp Lys Leu Ser Ser Ser Gly Ala Ser Trp His Thr Glu	
50	60

21062W0.ST25.txt

Trp Thr Trp Ser Gly Gly Glu Gly Thr Val Lys Ser Tyr Ser Asn Ser
 65 70 75 80
 Gly Val Thr Phe Asn Lys Lys Leu Val Ser Asp Val Ser Ser Ile Pro
 85 90 95
 Thr Ser Val Glu Trp Lys Gln Asp Asn Thr Asn Val Asn Ala Asp Val
 100 105 110
 Ala Tyr Asp Leu Phe Thr Ala Ala Asn Val Asp His Ala Thr Ser Ser
 115 120 125
 Gly Asp Tyr Glu Leu Met Ile Trp Leu Ala Arg Tyr Gly Asn Ile Gln
 130 135 140
 Pro Ile Gly Lys Gln Ile Ala Thr Ala Thr Val Gly Gly Lys Ser Trp
 145 150 155 160
 Glu Val Trp Tyr Gly Ser Thr Thr Gln Ala Gly Ala Glu Gln Arg Thr
 165 170 175
 Tyr Ser Phe Val Ser Glu Ser Pro Ile Asn Ser Tyr Ser Gly Asp Ile
 180 185 190
 Asn Ala Phe Phe Ser Tyr Leu Thr Gln Asn Gln Gly Phe Pro Ala Ser
 195 200 205
 Ser Gln Tyr Leu Ile Asn Leu Gln Phe Gly Thr Glu Ala Phe Thr Gly
 210 215 220
 Gly Pro Ala Thr Phe Thr Val Asp Asn Trp Thr Ala Ser Val Asn
 225 230 235

<210> 16

<211> 2659

<212> DNA

<213> *Aspergillus niger*

<400> 16

gatgatggtg atgttgaggg tattgatggt agttagggac tcgggagaga agactgttca 60
 agaggaagga ggggggaaat gctgcagtgc cagttggcaa tgccatcatc tgagggctta 120
 tgcggggtcc ggagcaattc cgcactggcg tcattgttgg attcgtcatc attcttgagg 180
 tcactgactc actcactatg caactggaac aactggatat gttgacttat gctgtgccaa 240
 ggctggccac gtggtggggg gattatccca gatactaaat tacatctata ttgactatct 300
 aactcatgtc tgcacactaa gtcggagtag taacctataa atacctgggt catgcttggc 360
 tctgattggt tcgtcatttg gggccgatac tagtagtgag cttggcaaga tcggcagcat 420
 ttgggtctca atatttatat cagtctattc caatgaacaa cgccctgaac cgcactgagt 480
 aacaatgcca tacagcttgt tggatatact actttttatc attggatttt ccaatggttg 540
 ataagcacca ttagcgacta attaacaact ggtgagctca tgtaagagag cttctgactg 600
 caacacttga ctatgaccaa ctaacttatt tatagtttag atcattatca ggaataacat 660
 gcttcgtagg cggttatata ccataagcaa agctgctaca ctccaactat gcgttgctac 720
 gagactttgc tgctgaccgc ttgaacatca gcatccggtt taccggaacg caagcagtg 780
 ggcgctatca agctccatta atctcacagc ttagactact tccgagggca gtgcaagctt 840

21062W0.ST25.txt

caagcaccaa	cttctccaac	aatgatcgat	cgatcgccgc	gaaacatgcc	cgatgtggcc	900
cgctaagctg	ccgtcagact	aggccacccg	ctaagccact	cgactgcgac	cgcaatgcgg	960
cattaggtga	acatttagca	ggattgtggt	gaagaaaatg	ctcccattcc	cactcattcc	1020
tgttgggggc	gaaaccttgt	tcccccttaa	tcaacagaaa	caaggaaata	aaaccgagga	1080
gttatccgac	caggttgaca	atttctcaca	cagtaccctc	gaccagtcca	ttgagtcttg	1140
attatcatca	tgaaccggta	cctgtctatc	cccctgtgct	tggcgagcgc	tgttcctgcg	1200
ctagcgggtc	tggcggcccg	cgaccccatc	ggtaccatct	gcaccaagaa	caacatcatc	1260
accaccgacg	acttcatect	ctacaataat	ctctggggcg	aggactatgc	cacttccggc	1320
tccgagtgtg	cctacctcga	ctatgacagc	ggcaactcca	tctcatggca	gacgtcgtgg	1380
acctgggctg	gcggcgacga	ctatgtcaag	agctatccca	atgccgtcct	gaacgtcggg	1440
gccaa gcagc	tgagcacgat	caccagtatt	ccctcgacct	ggaaatggag	gtgagccaat	1500
ctttcctgtc	acgaatgaag	ctattgtcta	atattcttag	ctacacgggc	aacgacctcg	1560
tcgccgacgt	ctcctacgat	gccttctctc	ccaccactga	ttccacaacc	gccactcacg	1620
agtacgaaat	catgatctgg	ctcgccgcct	acggtggcat	tgagcccatc	ggcaactcgg	1680
acggtcctat	cgccagcccc	actatcgggt	gctacacctg	ggatctctac	aagggtccca	1740
acgactggac	agtgtatagt	ttcgtggccc	gcgagactat	cacggacttc	tcggccgatg	1800
tcttgaggtt	ctttacttat	ttggttgata	acgaggggtg	ctcttcgagc	ttgtatttgc	1860
agaccctggg	tgcgggcact	gaaccgaaga	ctggctctga	tgcttggttc	acggtttcgc	1920
cgtatactgt	tagcatcaat	acttagaggg	ttcaagtagt	gtgacgttag	cagtgagtgg	1980
tttagttggg	gagaagtggg	gaagatatta	gagtataaaa	ttgaccaagc	ctgtgatatt	2040
ttttgtgaca	accagacggg	aattagtttc	gtcgataaatt	aatattggaa	atcttaaggg	2100
cgcaacgcta	tccataactc	ggttatttca	aatggtagaa	gtctggagaa	cgtttgtggc	2160
tgcggagtcc	ccatggcctg	ataggctaca	ttcctcgtcc	tatggagcag	tgctcttttt	2220
atctaaatga	ccgacgagaa	aaatttagat	aattaggaat	gggcatcatc	tgcgtcataa	2280
gatgatgtga	aaaaaatgaa	aaagaaatgt	ttgaactccg	atagggggaa	tcgaaccccc	2340
ggctgccgtg	tgagagacgg	cgatgttagc	cactacacca	tatcggatta	gttgttattc	2400
tcgaccggaa	tatacactta	ataggcaaag	accaactgcc	aatcacgtgc	ctcccgatcg	2460
cattgtctcta	aacaacttcc	tcacctttcc	tccatagtaa	gcgtacttga	tttaccacc	2520
ccatcatcat	caagttactt	agagtgggtt	tcagataaaa	ttgttttatg	ggcttctacc	2580
tgattttatcc	cagcgtgaga	attacaatgg	caaagggtta	agtccactat	ggtgtggtag	2640
tagcatttca	tgtttctac					2659

<210> 17

<211> 747

<212> DNA

<213> Aspergillus niger

<220>

<221> CDS

<222> (1)..(747)

<400> 17

atg aac cgg tac ctg tct atc ccc ctg tgc ttg gcg agc gct gtt cct

48

21062W0.ST25.txt

Met	Asn	Arg	Tyr	Leu	Ser	Ile	Pro	Leu	Cys	Leu	Ala	Ser	Ala	Val	Pro		
1				5					10					15			
gcg	cta	gcg	ggt	ctg	gcg	gcc	cgc	gac	ccc	atc	ggt	acc	atc	tgc	acc		96
Ala	Leu	Ala	Gly	Leu	Ala	Ala	Arg	Asp	Pro	Ile	Gly	Thr	Ile	Cys	Thr		
			20					25					30				
aag	aac	aac	atc	atc	acc	acc	gac	gac	ttc	atc	ctc	tac	aat	aat	ctc		144
Lys	Asn	Asn	Ile	Ile	Thr	Thr	Asp	Asp	Phe	Ile	Leu	Tyr	Asn	Asn	Leu		
		35					40				45						
tgg	ggc	gag	gac	tat	gcc	act	tcc	ggc	tcc	gag	tgt	acc	tac	ctc	gac		192
Trp	Gly	Glu	Asp	Tyr	Ala	Thr	Ser	Gly	Ser	Glu	Cys	Thr	Tyr	Leu	Asp		
	50					55				60							
tat	gac	agc	ggc	aac	tcc	atc	tca	tgg	cag	acg	tcg	tgg	acc	tgg	gct		240
Tyr	Asp	Ser	Gly	Asn	Ser	Ile	Ser	Trp	Gln	Thr	Ser	Trp	Thr	Trp	Ala		
65				70				75						80			
ggc	ggc	gac	gac	tat	gtc	aag	agc	tat	ccc	aat	gcc	gtc	ctg	aac	gtc		288
Gly	Gly	Asp	Asp	Tyr	Val	Lys	Ser	Tyr	Pro	Asn	Ala	Val	Leu	Asn	Val		
			85					90				95					
ggt	gcc	aag	cag	ctg	agc	acg	atc	acc	agt	att	ccc	tcg	acc	tgg	aaa		336
Gly	Ala	Lys	Gln	Leu	Ser	Thr	Ile	Thr	Ser	Ile	Pro	Ser	Thr	Trp	Lys		
		100					105				110						
tgg	agc	tac	acg	ggc	aac	gac	ctc	gtc	gcc	gac	gtc	tcc	tac	gat	gcc		384
Trp	Ser	Tyr	Thr	Gly	Asn	Asp	Leu	Val	Ala	Asp	Val	Ser	Tyr	Asp	Ala		
		115					120				125						
ttc	ctc	tcc	acc	act	gat	tcc	aca	acc	gcc	act	cac	gag	tac	gaa	atc		432
Phe	Leu	Ser	Thr	Thr	Asp	Ser	Thr	Thr	Ala	Thr	His	Glu	Tyr	Glu	Ile		
		130				135					140						
atg	atc	tgg	ctc	gcc	gcc	tac	ggt	ggc	att	gag	ccc	atc	ggc	aac	tcg		480
Met	Ile	Trp	Leu	Ala	Ala	Tyr	Gly	Gly	Ile	Glu	Pro	Ile	Gly	Asn	Ser		
145				150				155				160					
gac	ggt	cct	atc	gcc	agc	ccc	act	atc	ggt	ggc	tac	acc	tgg	gat	ctc		528
Asp	Gly	Pro	Ile	Ala	Ser	Pro	Thr	Ile	Gly	Gly	Tyr	Thr	Trp	Asp	Leu		
			165					170				175					
tac	aag	ggt	ccc	aac	gac	tgg	aca	gtg	tat	agt	ttc	gtg	gcc	cgc	gag		576
Tyr	Lys	Gly	Pro	Asn	Asp	Trp	Thr	Val	Tyr	Ser	Phe	Val	Ala	Arg	Glu		
			180					185				190					
act	atc	acg	gac	ttc	tcg	gcc	gat	gtc	ttg	gag	ttc	ttt	act	tat	ttg		624
Thr	Ile	Thr	Asp	Phe	Ser	Ala	Asp	Val	Leu	Glu	Phe	Phe	Thr	Tyr	Leu		
		195					200				205						
gtt	gat	aac	gag	ggt	gtc	tct	tcg	agc	ttg	tat	ttg	cag	acc	ctg	ggt		672
Val	Asp	Asn	Glu	Gly	Val	Ser	Ser	Ser	Leu	Tyr	Leu	Gln	Thr	Leu	Gly		
		210				215					220						
gcg	ggc	act	gaa	ccg	aag	act	ggc	tct	gat	gct	tgg	ttc	acg	gtt	tcg		720
Ala	Gly	Thr	Glu	Pro	Lys	Thr	Gly	Ser	Asp	Ala	Trp	Phe	Thr	Val	Ser		
225					230					235					240		

21062WO.ST25.txt

ccg tat act gtt agc atc aat act tag
 Pro Tyr Thr Val Ser Ile Asn Thr
 245

747

<210> 18

<211> 248

<212> PRT

<213> Aspergillus niger

<400> 18

Met Asn Arg Tyr Leu Ser Ile Pro Leu Cys Leu Ala Ser Ala Val Pro
 1 5 10 15
 Ala Leu Ala Gly Leu Ala Ala Arg Asp Pro Ile Gly Thr Ile Cys Thr
 20 25 30
 Lys Asn Asn Ile Ile Thr Thr Asp Asp Phe Ile Leu Tyr Asn Asn Leu
 35 40 45
 Trp Gly Glu Asp Tyr Ala Thr Ser Gly Ser Glu Cys Thr Tyr Leu Asp
 50 55 60
 Tyr Asp Ser Gly Asn Ser Ile Ser Trp Gln Thr Ser Trp Thr Trp Ala
 65 70 75 80
 Gly Gly Asp Asp Tyr Val Lys Ser Tyr Pro Asn Ala Val Leu Asn Val
 85 90 95
 Gly Ala Lys Gln Leu Ser Thr Ile Thr Ser Ile Pro Ser Thr Trp Lys
 100 105 110
 Trp Ser Tyr Thr Gly Asn Asp Leu Val Ala Asp Val Ser Tyr Asp Ala
 115 120 125
 Phe Leu Ser Thr Thr Asp Ser Thr Thr Ala Thr His Glu Tyr Glu Ile
 130 135 140
 Met Ile Trp Leu Ala Ala Tyr Gly Gly Ile Glu Pro Ile Gly Asn Ser
 145 150 155 160
 Asp Gly Pro Ile Ala Ser Pro Thr Ile Gly Gly Tyr Thr Trp Asp Leu
 165 170 175
 Tyr Lys Gly Pro Asn Asp Trp Thr Val Tyr Ser Phe Val Ala Arg Glu
 180 185 190
 Thr Ile Thr Asp Phe Ser Ala Asp Val Leu Glu Phe Phe Thr Tyr Leu
 195 200 205
 Val Asp Asn Glu Gly Val Ser Ser Ser Leu Tyr Leu Gln Thr Leu Gly
 210 215 220
 Ala Gly Thr Glu Pro Lys Thr Gly Ser Asp Ala Trp Phe Thr Val Ser
 225 230 235 240
 Pro Tyr Thr Val Ser Ile Asn Thr
 245

21062WO.ST25.txt

<210> 19

<211> 9103

<212> DNA

<213> *Aspergillus niger*

<400> 19

accaaaatgg	tggcggacgc	tgcgatgatg	gctaccaagc	atggacgggc	ggatcggacc	60
gcctgagacc	ttgggtgtgg	aggaaactgta	gtcttggtga	tatggacgaa	gcacccccat	120
tacaggataa	tcatccagcg	ccaaaaggta	gaactcctgt	cacgctccga	gttaattctc	180
ctgggtgcat	cgtaagaaag	aatggcctgg	ggtcgagtga	tgcgggcgga	ggcgggtgggc	240
ggccagtcaa	ggcgcgaggc	cggggaaggc	agatgactcg	gggcgagcat	ctgcacggca	300
cctgcagtca	acttgcatat	atagaccgca	gtagcgcagt	caggggcact	tttggaatca	360
taggctgcag	gccaaacatc	acgctttccc	gacattattc	agcgcctaata	cccgcagcac	420
catgcttgag	gaaccttggc	cctccttttt	tttttttttg	cccggaatgg	acgcgtcaaa	480
taactccttc	tgccaaggga	tgcccagcgg	aaccaggcgg	gcaccgcaca	cattcttctc	540
gcgttccaaa	cctgagattg	ttgtctagct	aaatcatcca	aacatggcgg	agttcaacca	600
acaatccggc	ttccggggcg	accatctaata	cagactctgt	tcaagtccgt	aacgctgtac	660
aacttctgga	ccatgcaggt	gattgctgct	gatgccgagg	ctagtgatgg	ctaggttggt	720
tcacgggcag	accacgaccc	aagggttagcg	gccgatcagt	cgggatgggt	taattaagga	780
agaagactgt	ctcctctgcg	gaaagcttgg	tgcactgatt	ccttaccac	agtgtaggat	840
atthttatct	cgatcgggaa	gaaagagacg	ggaagactga	gacatatcgg	accgtggact	900
tggctatcta	ctagtagcat	gtagtctact	acggaaatcg	attctacgag	taggagaaga	960
ttcatgttgt	agagcatccg	taaccatttc	ttttggtggg	aaggtaaaaa	gttaactagt	1020
ggttaatttc	tggtttgccc	ctggatgata	ctccctcagc	tcactcgcct	ctcttcactc	1080
gccggggccca	cccgtccatg	caggaccaac	cttttccttc	ccgcctgagt	caaggcatct	1140
accaaccttg	aatccaaggg	gactgactga	cgactcattc	tacagctgaa	cagcatgcgg	1200
aaacgtatat	tccgccgaga	gtatatacgt	atagcgggct	aatccatgga	tttaactagct	1260
tttctgccgt	ggaatccata	aatctgacgc	agcaagtga	gcttatcgag	atgaggatgc	1320
cacggccgcc	caggagcacc	ggcattactg	cccagactgg	cgtcggggaa	acaatggggc	1380
acgaggaatg	tctagacctg	ctcaatcagg	cagcctccag	tactggaatc	acaacggaaa	1440
gttgggtgcg	ggagcccga	actcctcccc	atcgagacat	accgtatggc	gggtcaaacc	1500
cagcatggca	gagggtttgt	cccgtccaca	gggtcgccat	ccataccaat	gcctgagtct	1560
agtacgtagc	gaccagctag	ctgccctggc	gctgtcaaac	ggccaaattc	cagcaacaag	1620
ccatgccgga	cgatcggctc	tgcagccatg	gccgagagat	agaaatgccc	agttgtctac	1680
gatctgggcc	acggcgaata	tgttgatcaa	aaccctggga	ggtatattag	cctaccggat	1740
ccgctgatcg	actggggatt	ttgatctttt	cctttggctc	tttttctctc	tcctcctgcc	1800
acctccccct	ctccttttcc	gaatccctgg	gcctgtgctt	ttcagatcgg	atcagatttc	1860
gttcgttttg	tttgtcgtct	acgcacccct	gggaggtagg	ctatccctcg	ttcagtcact	1920
agtcacgggt	agtagagagt	cgtgcgactc	tttgcaattc	gctctcttcg	gtcgttcttt	1980
gaatatccgg	cagcctagaa	tatctggccc	accagcattg	gggagtagga	ctaatagtag	2040
tgtgtgtgcg	tgtatacaac	tacactcgaa	cttcggtagg	taattagtcc	atggcagaca	2100
ctctttcccg	cactgttttc	cccaccacca	ttcaccagca	ccaccacaca	cacaagctct	2160
tcgtggccca	ccatgcatga	ctccagccgg	ttcagcacag	tgcaatgctg	ctgctactac	2220
ctgcactatt	atatcaagcg	accgccgcga	gaacgatggc	caccaatgag	taccgactaa	2280

21062W0.ST25.txt

ccccgtcta	gtgctttatc	gattcgtttc	atcccctgac	ccaggcgttt	caagcagcgg	2340
tcggcagacg	acgaccaccg	cagctaattg	gtgtgagtgg	agcgatcggg	gcgcgtgatt	2400
ggatcttgag	cagcttttct	tgcagacgga	cctccaccat	ctggcgatag	tcaagaagta	2460
atctccagcg	acggccaagg	catcgagtc	cagccagcac	caatcgttgg	gtcaggaacg	2520
gggcaagttt	gtgcctggat	cctttggtcc	ctcccgcgg	tgccagcaaa	tcttcccccc	2580
tccactttaa	cccaagtgtg	ataaccttca	tttacatcgg	ccgcttcctt	catcccagct	2640
gtgtgatttc	tttttattct	cccctcctcc	attccttttt	ttccccgtca	gttgtgtctt	2700
tgtcgtggga	aaaggcagtt	tccgagtgtt	actgtcttgc	tgtgtgttgt	gttttttgtgt	2760
gtttttctta	cattcattcc	ttttctacta	cttgtctccc	ctggcggtcg	attcgcgccc	2820
ctccatacat	tccttgctct	tcccttattt	actcctccaa	ctacaaccac	ttttatccag	2880
ctcaccgtct	gaccccgctc	tgtgacaaac	cacttctggt	ggtgctcaca	cgatcgtgtc	2940
tggatatttc	ttcgggtactg	gttactgaag	tgttggctgc	tcgtcagcat	ttctcactat	3000
attactgtcc	cttctcgtc	ttcagttctc	tgccggctca	gtgaacctcg	ctgtcttgac	3060
gatggcgtct	ctccacactt	ccatcatggt	gtttgtgtcg	tacgttttcc	tgctgctgag	3120
tgtgccagt	gccttggtca	atgctgtccc	tacgtccggc	tctgtcgatg	tgacatccca	3180
tctggcgggtg	cgcgcgctcg	actactgggt	gggaacaatc	aagcgtcagg	gctctgtgcc	3240
ctttggcaac	accaccagca	gctatgcaat	ctaccgcaac	gttaaggact	acggtgccaa	3300
gggtgagtca	gacctcctta	catgactgggt	atctgccgct	gacccagaca	ggtgacggtt	3360
ctaccgatga	tactgatgct	atcaacaagg	ccatctcctc	tgggtggccgc	tgcggttccg	3420
gttgcgactc	gtctaccacc	acccggcccc	tgggtgactt	cccagcgggt	acctacgtgg	3480
tgtccaagcc	cattgtgcag	tactactaca	cccagctgggt	cgggtgatgcc	aacgacttgc	3540
ctgtcctcaa	ggctgcttcg	agcttcagcg	gcatggccgt	cattgatgcc	gacccctacg	3600
acagcaatgg	ctccaactgg	tacaccaacc	agaacaactt	cttccgtgct	gtccgcaact	3660
ttgtcatcga	tctgacggcc	atgcctgcat	cctcgggcgc	gggtattcac	tgccaggtcg	3720
gccaggccac	cagtctgcag	aacattcgct	tcgagatgggt	caaggggtggc	agcaacaaca	3780
agcagcaggg	tatcttcatg	gacaacggat	cgggtgggtt	tatgtctgac	ctcaccttca	3840
acggcggtaa	ctacggtgcc	ttcttcggta	accagcagtt	caccacccgc	aacctgacct	3900
tcaacgactg	caacacggcc	atctacatga	actggaactg	ggcctggact	ttcaagtccg	3960
tgtcgatcaa	caactgccag	gtcggcttga	acatgtccgc	atcgccgtcc	aaccagaccg	4020
tcggatctgt	cctgatgcta	gacagacat	tcactagcac	tcctactgggt	gtggtgactg	4080
ccttcagcag	cgacagcatt	cctactgggt	gtggtgcatt	ggctcctgaa	aatgtcgact	4140
tctccggcgc	tacgggtggcc	attgctagt	tgaacggcga	caccattctc	aagggcggct	4200
ctgttgtctc	gaactacgtg	cagggcaatg	tctacacgcc	gggcagcacc	tccaacagca	4260
ccgtcagcaa	gcgctcccct	gccgttgagg	ccgaggatgt	tcatgtggtc	acccagacca	4320
ttaccgagac	cattacagcc	tgcggtgccg	actacgccac	tgaaaccgct	gcgggttcgt	4380
tcggcacggc	ctctgaggcc	ggagcagagg	gttcgaccac	taccgcggcc	ggtggtgtgc	4440
tgcttttgcc	atccatgtct	gtgcagtacc	cgtccgccgg	tagctacagc	tgggaaggat	4500
ttgacagcga	aacctctcag	gcaaccacgg	ccaccaacga	ggcttctgtc	cctgccgccc	4560
gaacttcttg	cgggtcatct	cctgccgtgc	ccacttccca	gtcggccgggt	gcctctgttg	4620
ttgccagggg	ctctggctct	gcgtccactg	cctccagcac	ctcgtcttcc	ggctcctact	4680
ttggtacttc	cagcaactcc	accggccagt	gctcgaccaa	gactgtcacc	aagacccgta	4740
tcgagaccac	tattccttcc	caggccatgg	ctagctcgct	tctgactagc	tctggtagt	4800
tgttcgagaa	gtccaagcct	ttgtacgcct	cctatgccgc	ttcttccttc	gtcagcgta	4860
agtcggccgg	tgccaagggt	gacggtacca	ccgatgacac	ggccatcatc	cagagcgttc	4920

21062W0.ST25.txt

tgaacaacgc	gaccgaggac	cagatcgtgt	acttcgatca	cgggtgcctac	atcatcactg	4980
ataccatcaa	ggtccccaag	aacatcaaga	tcactggtga	ggtgtggccc	atgctcatgg	5040
cgtatggcga	caagttctcc	gaccagaaga	accccatccc	tatgctccag	gtcgggtgaag	5100
ccggcgagtc	tggctcgggt	gagatcagcg	acctcgttct	ggagaccaag	ggacccgctc	5160
ctggtgccat	cctggtgcag	tggaaacctc	aggccgcttc	ccagggctct	gttggtatgt	5220
gggatgtcca	tggccgcatt	ggtggctctg	ccggcactga	acttcagtcc	gatacctgcg	5280
cgaagacccc	cagccagaca	accacggcca	acgagaagtg	cattggcgcc	ttcctcctct	5340
tccacgccac	ccagaacgcc	agcgtctacc	tggagaacac	ctggttctgg	actgctgacc	5400
acgaccttga	tgcgtccggc	cacgggcaga	tcaacattta	caacggtcgc	ggtgtcctcc	5460
tcgagagtga	gggaccagcc	tggctgtatg	gcacggctgc	cgagcacaac	cagctgtaca	5520
actaccaggt	ctccaatgcc	agaacgttt	tcatgggcct	gatccagacc	gagactccct	5580
actaccaggc	caacccaac	gctctgactc	ccttcacccc	gcagacctcc	tggaatgacc	5640
ccgacttctc	gtactgcacc	acggactcct	gccgcaaggc	gtggggctct	cgtgtcatca	5700
acagctccga	gacctacatc	tacggtgcgg	gtctgtacag	cttcttcgac	aactactcgc	5760
agacatgtct	ggccacggag	tcctgccagg	agaacattgt	tgaggtggac	tgctccgacg	5820
tgcacatcta	cgggctgagc	accaaggcca	gcaccaacat	gatcacttcg	tccagcggcg	5880
atggccttgt	cccgaagag	ggcaacgaga	gcacatactg	ctcgacgac	gctttgttcg	5940
agcagtcctc	ctctatttta	tttgtgcttt	tacatcagag	acaatacata	catacagtga	6000
taaagcagtt	gtttgtttcg	atgcccgtgg	aggaaaggga	acttagatgc	atttgggatg	6060
gttgattgtt	tgttggggaa	cacatacgga	tataccactc	tgttggttaca	ttttatatatt	6120
ttccctatcc	attcatttta	ttcttcgatt	cttatcccat	tcgttccatc	catttggtgt	6180
atacaggccc	tggtaatagt	cgaagtggta	gaggggggtt	agcagtagag	gtgctccgca	6240
tgtggctata	tagtagttgg	aataaaaagt	caatgaccag	attgtttcta	ctagatgcta	6300
gcatacata	aatgaagttc	tatttagctc	atgctcttca	tcccattgac	attgatactc	6360
tacatacgta	cccattccgc	catcaccatt	gttaccctct	gattatcagc	acccttaaac	6420
acttagttac	caaacagcaa	tagttgactt	gattcatcca	agtaatacgc	tctatcttct	6480
tctcccttca	cctcaatctc	ttgatataat	tctctagtac	caatactcac	aacctacccg	6540
cacccaaaag	ccacaacgag	aacaacaacc	cacacgacca	aattacctga	tcgtatatga	6600
caagatagac	aaagaaacaa	aaaacagtg	ttgattcaat	tcaagtgaac	gttcttgctc	6660
ctttctaata	cataatccat	tttttctttt	ccctttcaat	tgatctgcat	tatgcattgt	6720
cttgtttcat	agtattcatt	atcttattgt	atgcacaata	gccattgtct	agtcagttat	6780
aggtaaacta	atttggacag	agagtggatg	tgaatgtgta	gaaagtgggg	tattgttgaa	6840
tagaatagaa	gggatataga	tgtaataggt	gaatgaaatg	tattgattta	ttgatgaaat	6900
ggttggtttt	tgttcttggt	tcaattgata	attggtgtgt	ttggtgagaa	gatgaatttg	6960
ccctgtcttg	gtctgggggt	aaggggactt	gggtgggcta	gtcggttggg	ttagtattta	7020
gtttgtttac	tgctttgtga	tggatgatat	tctgggttag	gtttcttttg	gttggttttg	7080
gtagtgtatt	tattgattcg	attcgataat	ctcttgcgag	tatttggttag	gttagtctgt	7140
tgatctatgt	aaggatgggt	tggccagtga	cattggaagt	atgcttttct	attgtgagag	7200
agattatggt	gatgatgatg	taaaggattt	gtaggatatg	tgtgctgctg	atatttccat	7260
gtattatttg	ttgcatcgct	gagattctac	tgatgagtgt	gtcgtgggag	aagcggatga	7320
acgcctgagg	gacaaaagg	tatcaatggg	tgagaggtag	cagaccttgc	cctgctattc	7380
ttcgcaataa	tgtttgtgtg	tgtggtgttt	ggtgaacggt	cgttgacttt	tcgtaggggg	7440
ctttttttca	ttggagggtg	gggctatatt	gactccgcgg	tcgagctggg	aaggagttgg	7500
ggctatttgc	ttgactggag	gttaatgtag	ctgtgattat	ctccgagtgt	tagcacaata	7560

21062W0.ST25.txt

tgatgtgtcc	tatattccca	acgattggct	ttgtggcagc	agctggatac	gtatctgcat	7620
tcctctggac	ttgtaagaat	atcaagccag	gaatttttcc	ggcccctctc	ctccgatagt	7680
ggagtgttat	aactggcaga	agacaatcgg	tcctgatggg	agaagagtac	caacgggtcg	7740
ttatgatagt	ccagtttagt	tgaggctcat	atgaaggcca	gtgtatgggg	aattgcagtt	7800
ctcaatgtcg	tccggttcca	attggggcac	tggcggattg	gattgatcaa	atgtccctca	7860
agacatctgt	gggtcatcgt	agtctgccgc	tgtttcctcc	agggactggc	aacagtagag	7920
ggaatcattt	ctgctcaatt	atgaatgtca	tttagcagga	ccaaaagata	gcactgttga	7980
atgacgagca	gtacgctaga	tgaagaggca	gacgagaaac	ctttcacact	tggcactttg	8040
caagatgaag	agagacatcc	aagacttgct	aataattttg	cttgtcagca	aggggaaggc	8100
gggggggact	gttttgaata	atcctcatcc	tggttgagag	agctccagta	ggagatatgc	8160
acagctcctt	tccctgctag	gaatagtggg	ggaagacctg	aggactgaaa	ggtatgagat	8220
gaatgaacca	ggcaagacat	gcgggtggac	aacgtccgag	tcgagcagaa	caccatcctc	8280
aatgctgcca	ctgctataac	gccagggctg	atgcttgctg	ggagtaaatt	gagtcccagc	8340
tggtgccatg	gcttgattgg	cgccgttccg	gatggatccc	gaaccttggc	ccgagtctaa	8400
ttggggagta	gatccagtta	tgacggggcc	atttcggaaa	tgaaggaagg	aaagaggaaa	8460
gccagcagat	agtctagata	gtagataccg	aacagagaaa	aaaaaccaa	gccggagggt	8520
caccgagaaa	aggagagcgc	atggttggca	gtcaacaggc	gcggcgtaa	tgaagctgaa	8580
acccgggcaa	tgccaagcat	cccgggtcaat	ccgccggccc	tgaattcaac	tccacccgcc	8640
gtttcaacgc	ccctcaagat	gaaacctctg	cgctgctgcc	agcgcttgag	ccagtccagt	8700
cagtctccag	tccctgggtg	agtgactccc	ctgtcattat	tatcttcagg	gtttttggag	8760
ggccctctcg	gtgatgctag	aaccgtttca	cgatcgatca	tgatggagag	ctttgttcag	8820
gcccagaacc	cgccggctgg	cttcaacgcc	tatatagaac	ggcctctcgt	cgccctcact	8880
ctgctcttgt	tttctttcct	tcagcacaca	ggcttgctta	agtttcaatt	catcctttct	8940
ctttcccaca	ttctctttcc	ccggtctgtg	accgtcacat	tctttccaac	caaaactcat	9000
caacttccat	tccttgaaga	ataaccatct	ggtataccgg	tataccagac	tctttcacta	9060
tctttctatt	ctcctagaaa	aaaaaatctt	tcaatccacc	aaa		9103

<210> 20

<211> 2826

<212> DNA

<213> Aspergillus niger

<220>

<221> CDS

<222> (1)..(2826)

<400> 20

atg	ttg	ttt	gtg	tcg	tac	gtt	ttc	ctg	ctg	ctg	agt	gtg	cca	gtg	gcc	48
Met	Leu	Phe	Val	Ser	Tyr	Val	Phe	Leu	Leu	Leu	Ser	Val	Pro	Val	Ala	
1			5					10					15			
ttg	gtc	aat	gct	gtc	cct	acg	tcc	ggc	tct	gtc	gat	gtg	aca	tcc	cat	96
Leu	Val	Asn	Ala	Val	Pro	Thr	Ser	Gly	Ser	Val	Asp	Val	Thr	Ser	His	
		20						25					30			

21062WO.ST25.txt

ctg gcg gtg cgc gcg tcg gac tac tgg gtg gga aca atc aag cgt cag	144
Leu Ala Val Arg Ala Ser Asp Tyr Trp Val Gly Thr Ile Lys Arg Gln	
35 40 45	
ggc tct gtg ccc ttt ggc aac acc acc agc agc tat gca atc tac cgc	192
Gly Ser Val Pro Phe Gly Asn Thr Thr Ser Ser Tyr Ala Ile Tyr Arg	
50 55 60	
aac gtt aag gac tac ggt gcc aag ggt gac ggt tct acc gat gat act	240
Asn Val Lys Asp Tyr Gly Ala Lys Gly Asp Gly Ser Thr Asp Asp Thr	
65 70 75 80	
gat gct atc aac aag gcc atc tcc tct ggt ggc cgc tgc ggt tcc ggt	288
Asp Ala Ile Asn Lys Ala Ile Ser Ser Gly Gly Arg Cys Gly Ser Gly	
85 90 95	
tgc gac tcg tct acc acc acc ccg gcc ctg gtg tac ttc cca gcg ggt	336
Cys Asp Ser Ser Thr Thr Thr Pro Ala Leu Val Tyr Phe Pro Ala Gly	
100 105 110	
acc tac gtg gtg tcc aag ccc att gtg cag tac tac tac acc cag ctg	384
Thr Tyr Val Val Ser Lys Pro Ile Val Gln Tyr Tyr Tyr Thr Gln Leu	
115 120 125	
gtc ggt gat gcc aac gac ttg cct gtc ctc aag gct gct tcg agc ttc	432
Val Gly Asp Ala Asn Asp Leu Pro Val Leu Lys Ala Ala Ser Ser Phe	
130 135 140	
agc ggc atg gcc gtc att gat gcc gac ccc tac gac agc aat ggc tcc	480
Ser Gly Met Ala Val Ile Asp Ala Asp Pro Tyr Asp Ser Asn Gly Ser	
145 150 155 160	
aac tgg tac acc aac cag aac aac ttc ttc cgt gct gtc cgc aac ttt	528
Asn Trp Tyr Thr Asn Gln Asn Asn Phe Phe Arg Ala Val Arg Asn Phe	
165 170 175	
gtc atc gat ctg acg gcc atg cct gca tcc tcg ggc gcg ggt att cac	576
Val Ile Asp Leu Thr Ala Met Pro Ala Ser Ser Gly Ala Gly Ile His	
180 185 190	
tgg cag gtc ggc cag gcc acc agt ctg cag aac att cgc ttc gag atg	624
Trp Gln Val Gly Gln Ala Thr Ser Leu Gln Asn Ile Arg Phe Glu Met	
195 200 205	
gtc aag ggt ggc agc aac aac aag cag cag ggt atc ttc atg gac aac	672
Val Lys Gly Gly Ser Asn Asn Lys Gln Gln Gly Ile Phe Met Asp Asn	
210 215 220	
gga tcg ggt ggt ttt atg tct gac ctc acc ttc aac ggc ggt aac tac	720
Gly Ser Gly Gly Phe Met Ser Asp Leu Thr Phe Asn Gly Gly Asn Tyr	
225 230 235 240	
ggt gcc ttc ttc ggt aac cag cag ttc acc acc cgc aac ctg acc ttc	768
Gly Ala Phe Phe Gly Asn Gln Gln Phe Thr Thr Arg Asn Leu Thr Phe	
245 250 255	
aac gac tgc aac acg gcc atc tac atg aac tgg aac tgg gcc tgg act	816
Asn Asp Cys Asn Thr Ala Ile Tyr Met Asn Trp Asn Trp Ala Trp Thr	

21062W0.ST25.txt

260	265	270	
ttc aag tcg gtg tcg atc aac aac tgc cag gtc ggc ttg aac atg tcc			864
Phe Lys Ser Val Ser Ile Asn Asn Cys Gln Val Gly Leu Asn Met Ser			
275	280	285	
gca tcg ccg tcc aac cag acc gtc gga tct gtc ctg atg cta gac agc			912
Ala Ser Pro Ser Asn Gln Thr Val Gly Ser Val Leu Met Leu Asp Ser			
290	295	300	
aca ttc act agc act cct act ggt gtg gtg act gcc ttc agc agc gac			960
Thr Phe Thr Ser Thr Pro Thr Gly Val Val Thr Ala Phe Ser Ser Asp			
305	310	315	320
agc att cct act ggt ggt ggt gca ttg gtc ctc gaa aat gtc gac ttc			1008
Ser Ile Pro Thr Gly Gly Gly Ala Leu Val Leu Glu Asn Val Asp Phe			
325	330	335	
tcc ggc gct acg gtg gcc att gct agt gtg aac ggc gac acc att ctc			1056
Ser Gly Ala Thr Val Ala Ile Ala Ser Val Asn Gly Asp Thr Ile Leu			
340	345	350	
aag ggc ggc tct gtt gtc tcg aac tac gtg cag ggc aat gtc tac acg			1104
Lys Gly Gly Ser Val Val Ser Asn Tyr Val Gln Gly Asn Val Tyr Thr			
355	360	365	
ccg ggc agc acc tcc aac agc acc gtc agc aag cgc tcc cct gcc gtt			1152
Pro Gly Ser Thr Ser Asn Ser Thr Val Ser Lys Arg Ser Pro Ala Val			
370	375	380	
gag gcc gag gat gtt cat gtg gtc acc cag acc att acc gag acc att			1200
Glu Ala Glu Asp Val His Val Val Thr Gln Thr Ile Thr Glu Thr Ile			
385	390	395	400
aca gcc tgc ggt gcc gac tac gcc act gaa acc gct gcg ggt tcg ttc			1248
Thr Ala Cys Gly Ala Asp Tyr Ala Thr Glu Thr Ala Ala Gly Ser Phe			
405	410	415	
ggc acg gcc tct gag gcc gga gca gag ggt tcg acc act acc gcg gcc			1296
Gly Thr Ala Ser Glu Ala Gly Ala Glu Gly Ser Thr Thr Thr Ala Ala			
420	425	430	
ggt ggt gtg ctg cct ttg cca tcc atg tct gtg cag tac ccg tcc gcc			1344
Gly Gly Val Leu Pro Leu Pro Ser Met Ser Val Gln Tyr Pro Ser Ala			
435	440	445	
ggt agc tac agc tgg gaa gga ttt gac agc gaa acc tct cag gca acc			1392
Gly Ser Tyr Ser Trp Glu Gly Phe Asp Ser Glu Thr Ser Gln Ala Thr			
450	455	460	
acg gcc acc aac gag gct tct gtc cct gcc gcc gga act tct ggc ggg			1440
Thr Ala Thr Asn Glu Ala Ser Val Pro Ala Ala Gly Thr Ser Gly Gly			
465	470	475	480
tca tct cct gcc gtg ccc act tcc cag tcg gcc ggt gcc tct gtt gtt			1488
Ser Ser Pro Ala Val Pro Thr Ser Gln Ser Ala Gly Ala Ser Val Val			
485	490	495	
gcc cag ggc tct ggc tct gcg tcc act gcc tcc agc acc tcg tct tcc			1536

Ala	Gln	Gly	Ser	Gly	Ser	Ala	Ser	Thr	Ala	Ser	Ser	Thr	Ser	Ser	Ser	
			500				505						510			
ggc	tcc	tac	ttt	ggt	act	tcc	agc	aac	tcc	acc	ggc	cag	tgc	tcg	acc	1584
Gly	Ser	Tyr	Phe	Gly	Thr	Ser	Ser	Asn	Ser	Thr	Gly	Gln	Cys	Ser	Thr	
			515				520						525			
aag	act	gtc	acc	aag	acc	cgt	atc	gag	acc	act	att	cct	tcc	cag	gcc	1632
Lys	Thr	Val	Thr	Lys	Thr	Arg	Ile	Glu	Thr	Thr	Ile	Pro	Ser	Gln	Ala	
			530				535						540			
atg	gct	agc	tcg	ctt	ctg	act	agc	tct	ggt	agt	gtg	ttc	gag	aag	tcc	1680
Met	Ala	Ser	Ser	Leu	Leu	Thr	Ser	Ser	Gly	Ser	Val	Phe	Glu	Lys	Ser	
			545				550						555			560
aag	cct	ttg	tac	gcc	tcc	tat	gcc	gct	tct	tcc	ttc	gtc	agc	gtc	aag	1728
Lys	Pro	Leu	Tyr	Ala	Ser	Tyr	Ala	Ala	Ser	Ser	Phe	Val	Ser	Val	Lys	
			565				570						575			
tcg	gcc	ggt	gcc	aag	ggt	gac	ggt	acc	acc	gat	gac	acg	gcc	atc	atc	1776
Ser	Ala	Gly	Ala	Lys	Gly	Asp	Gly	Thr	Thr	Asp	Asp	Thr	Ala	Ile	Ile	
			580				585						590			
cag	agc	gtt	ctg	aac	aac	gcg	acc	gag	gac	cag	atc	gtg	tac	ttc	gat	1824
Gln	Ser	Val	Leu	Asn	Asn	Ala	Thr	Glu	Asp	Gln	Ile	Val	Tyr	Phe	Asp	
			595				600						605			
cac	ggt	gcc	tac	atc	atc	act	gat	acc	atc	aag	gtc	ccc	aag	aac	atc	1872
His	Gly	Ala	Tyr	Ile	Ile	Thr	Asp	Thr	Ile	Lys	Val	Pro	Lys	Asn	Ile	
			610				615						620			
aag	atc	act	ggt	gag	gtg	tgg	ccc	atg	ctc	atg	gcg	tat	ggc	gac	aag	1920
Lys	Ile	Thr	Gly	Glu	Val	Trp	Pro	Met	Leu	Met	Ala	Tyr	Gly	Asp	Lys	
			625				630						635			640
ttc	tcc	gac	cag	aag	aac	ccc	atc	cct	atg	ctc	cag	gtc	ggt	gaa	gcc	1968
Phe	Ser	Asp	Gln	Lys	Asn	Pro	Ile	Pro	Met	Leu	Gln	Val	Gly	Glu	Ala	
			645				650						655			
ggc	gag	tct	ggc	tcg	gtt	gag	atc	agc	gac	ctc	gtt	ctg	gag	acc	aag	2016
Gly	Glu	Ser	Gly	Ser	Val	Glu	Ile	Ser	Asp	Leu	Val	Leu	Glu	Thr	Lys	
			660				665						670			
gga	ccc	gct	cct	ggt	gcc	atc	ctg	gtg	cag	tgg	aac	ctc	cag	gcc	gct	2064
Gly	Pro	Ala	Pro	Gly	Ala	Ile	Leu	Val	Gln	Trp	Asn	Leu	Gln	Ala	Ala	
			675				680						685			
tcc	cag	ggc	tct	gtt	ggt	atg	tgg	gat	gtc	cat	gcc	cgc	att	ggt	ggc	2112
Ser	Gln	Gly	Ser	Val	Gly	Met	Trp	Asp	Val	His	Ala	Arg	Ile	Gly	Gly	
			690				695						700			
tct	gcc	ggc	act	gaa	ctt	cag	tcc	gat	acc	tgc	gcg	aag	acc	ccc	agc	2160
Ser	Ala	Gly	Thr	Glu	Leu	Gln	Ser	Asp	Thr	Cys	Ala	Lys	Thr	Pro	Ser	
			705				710						715			720
cag	aca	acc	acg	gcc	aac	gag	aag	tgc	att	ggc	gcc	ttc	ctc	ctc	ttc	2208
Gln	Thr	Thr	Thr	Ala	Asn	Glu	Lys	Cys	Ile	Gly	Ala	Phe	Leu	Leu	Phe	

21062W0.ST25.txt

cac gcc acc cag aac gcc agc gtc tac ctg gag aac acc tgg ttc tgg	2256
His Ala Thr Gln Asn Ala Ser Val Tyr Leu Glu Asn Thr Trp Phe Trp	
740 745 750	
act gct gac cac gac ctt gat gcg tcc ggc cac ggg cag atc aac att	2304
Thr Ala Asp His Asp Leu Asp Ala Ser Gly His Gly Gln Ile Asn Ile	
755 760 765	
tac aac ggt cgc ggt gtc ctc ctc gag agt gag gga cca gcc tgg ctg	2352
Tyr Asn Gly Arg Gly Val Leu Leu Glu Ser Glu Gly Pro Ala Trp Leu	
770 775 780	
tat ggc acg gct gcc gag cac aac cag ctg tac aac tac cag gtc tcc	2400
Tyr Gly Thr Ala Ala Glu His Asn Gln Leu Tyr Asn Tyr Gln Val Ser	
785 790 795 800	
aat gcc aag aac gtt ttc atg ggc ctg atc cag acc gag act ccc tac	2448
Asn Ala Lys Asn Val Phe Met Gly Leu Ile Gln Thr Glu Thr Pro Tyr	
805 810 815	
tac cag gcc aac ccc aac gct ctg act ccc ttc acc ccg cag acc tcc	2496
Tyr Gln Ala Asn Pro Asn Ala Leu Thr Pro Phe Thr Pro Gln Thr Ser	
820 825 830	
tgg aat gac ccc gac ttc tcg tac tgc acc acg gac tcc tgc cgc aag	2544
Trp Asn Asp Pro Asp Phe Ser Tyr Cys Thr Thr Asp Ser Cys Arg Lys	
835 840 845	
gcg tgg ggt ctc cgt gtc atc aac agc tcc gag acc tac atc tac ggt	2592
Ala Trp Gly Leu Arg Val Ile Asn Ser Ser Glu Thr Tyr Ile Tyr Gly	
850 855 860	
gcg ggt ctg tac agc ttc ttc gac aac tac tcg cag aca tgt ctg gcc	2640
Ala Gly Leu Tyr Ser Phe Phe Asp Asn Tyr Ser Gln Thr Cys Leu Ala	
865 870 875 880	
acg gag tcc tgc cag gag aac att gtt gag gtg gac tgc tcc gac gtg	2688
Thr Glu Ser Cys Gln Glu Asn Ile Val Glu Val Asp Cys Ser Asp Val	
885 890 895	
cac atc tac ggg ctg agc acc aag gcc agc acc aac atg atc act tcg	2736
His Ile Tyr Gly Leu Ser Thr Lys Ala Ser Thr Asn Met Ile Thr Ser	
900 905 910	
tcc agc ggc gat ggc ctt gtc ccg caa gag ggc aac gag agc aca tac	2784
Ser Ser Gly Asp Gly Leu Val Pro Gln Glu Gly Asn Glu Ser Thr Tyr	
915 920 925	
tgc tcg acg atc gct ttg ttc gag cag tcc tcc tct att taa	2826
Cys Ser Thr Ile Ala Leu Phe Glu Gln Ser Ser Ser Ile	
930 935 940	

<210> 21

<211> 941

<212> PRT

21062W0.ST25.txt

<213> Aspergillus niger

<400> 21

```

Met Leu Phe Val Ser Tyr Val Phe Leu Leu Leu Ser Val Pro Val Ala
1           5           10           15
Leu Val Asn Ala Val Pro Thr Ser Gly Ser Val Asp Val Thr Ser His
           20           25           30
Leu Ala Val Arg Ala Ser Asp Tyr Trp Val Gly Thr Ile Lys Arg Gln
           35           40           45
Gly Ser Val Pro Phe Gly Asn Thr Thr Ser Ser Tyr Ala Ile Tyr Arg
           50           55           60
Asn Val Lys Asp Tyr Gly Ala Lys Gly Asp Gly Ser Thr Asp Asp Thr
65           70           75           80
Asp Ala Ile Asn Lys Ala Ile Ser Ser Gly Gly Arg Cys Gly Ser Gly
           85           90           95
Cys Asp Ser Ser Thr Thr Thr Pro Ala Leu Val Tyr Phe Pro Ala Gly
           100          105          110
Thr Tyr Val Val Ser Lys Pro Ile Val Gln Tyr Tyr Tyr Thr Gln Leu
           115          120          125
Val Gly Asp Ala Asn Asp Leu Pro Val Leu Lys Ala Ala Ser Ser Phe
           130          135          140
Ser Gly Met Ala Val Ile Asp Ala Asp Pro Tyr Asp Ser Asn Gly Ser
145          150          155          160
Asn Trp Tyr Thr Asn Gln Asn Asn Phe Phe Arg Ala Val Arg Asn Phe
           165          170          175
Val Ile Asp Leu Thr Ala Met Pro Ala Ser Ser Gly Ala Gly Ile His
           180          185          190
Trp Gln Val Gly Gln Ala Thr Ser Leu Gln Asn Ile Arg Phe Glu Met
           195          200          205
Val Lys Gly Gly Ser Asn Asn Lys Gln Gln Gly Ile Phe Met Asp Asn
           210          215          220
Gly Ser Gly Gly Phe Met Ser Asp Leu Thr Phe Asn Gly Gly Asn Tyr
225          230          235          240
Gly Ala Phe Phe Gly Asn Gln Gln Phe Thr Thr Arg Asn Leu Thr Phe
           245          250          255
Asn Asp Cys Asn Thr Ala Ile Tyr Met Asn Trp Asn Trp Ala Trp Thr
           260          265          270
Phe Lys Ser Val Ser Ile Asn Asn Cys Gln Val Gly Leu Asn Met Ser
           275          280          285
Ala Ser Pro Ser Asn Gln Thr Val Gly Ser Val Leu Met Leu Asp Ser
           290          295          300
Thr Phe Thr Ser Thr Pro Thr Gly Val Val Thr Ala Phe Ser Ser Asp
305          310          315          320
Ser Ile Pro Thr Gly Gly Gly Ala Leu Val Leu Glu Asn Val Asp Phe

```

21062W0.ST25.txt

325 330 335
 Ser Gly Ala Thr Val Ala Ile Ala Ser Val Asn Gly Asp Thr Ile Leu
 340 345 350
 Lys Gly Gly Ser Val Val Ser Asn Tyr Val Gln Gly Asn Val Tyr Thr
 355 360 365
 Pro Gly Ser Thr Ser Asn Ser Thr Val Ser Lys Arg Ser Pro Ala Val
 370 375 380
 Glu Ala Glu Asp Val His Val Val Thr Gln Thr Ile Thr Glu Thr Ile
 385 390 395 400
 Thr Ala Cys Gly Ala Asp Tyr Ala Thr Glu Thr Ala Ala Gly Ser Phe
 405 410 415
 Gly Thr Ala Ser Glu Ala Gly Ala Glu Gly Ser Thr Thr Thr Ala Ala
 420 425 430
 Gly Gly Val Leu Pro Leu Pro Ser Met Ser Val Gln Tyr Pro Ser Ala
 435 440 445
 Gly Ser Tyr Ser Trp Glu Gly Phe Asp Ser Glu Thr Ser Gln Ala Thr
 450 455 460
 Thr Ala Thr Asn Glu Ala Ser Val Pro Ala Ala Gly Thr Ser Gly Gly
 465 470 475 480
 Ser Ser Pro Ala Val Pro Thr Ser Gln Ser Ala Gly Ala Ser Val Val
 485 490 495
 Ala Gln Gly Ser Gly Ser Ala Ser Thr Ala Ser Ser Thr Ser Ser Ser
 500 505 510
 Gly Ser Tyr Phe Gly Thr Ser Ser Asn Ser Thr Gly Gln Cys Ser Thr
 515 520 525
 Lys Thr Val Thr Lys Thr Arg Ile Glu Thr Thr Ile Pro Ser Gln Ala
 530 535 540
 Met Ala Ser Ser Leu Leu Thr Ser Ser Gly Ser Val Phe Glu Lys Ser
 545 550 555 560
 Lys Pro Leu Tyr Ala Ser Tyr Ala Ala Ser Ser Phe Val Ser Val Lys
 565 570 575
 Ser Ala Gly Ala Lys Gly Asp Gly Thr Thr Asp Asp Thr Ala Ile Ile
 580 585 590
 Gln Ser Val Leu Asn Asn Ala Thr Glu Asp Gln Ile Val Tyr Phe Asp
 595 600 605
 His Gly Ala Tyr Ile Ile Thr Asp Thr Ile Lys Val Pro Lys Asn Ile
 610 615 620
 Lys Ile Thr Gly Glu Val Trp Pro Met Leu Met Ala Tyr Gly Asp Lys
 625 630 635 640
 Phe Ser Asp Gln Lys Asn Pro Ile Pro Met Leu Gln Val Gly Glu Ala
 645 650 655
 Gly Glu Ser Gly Ser Val Glu Ile Ser Asp Leu Val Leu Glu Thr Lys
 660 665 670
 Gly Pro Ala Pro Gly Ala Ile Leu Val Gln Trp Asn Leu Gln Ala Ala

21062WO.ST25.txt

```

        675                680                685
Ser Gln Gly Ser Val Gly Met Trp Asp Val His Ala Arg Ile Gly Gly
   690                695                700
Ser Ala Gly Thr Glu Leu Gln Ser Asp Thr Cys Ala Lys Thr Pro Ser
705                710                715                720
Gln Thr Thr Thr Ala Asn Glu Lys Cys Ile Gly Ala Phe Leu Leu Phe
   725                730                735
His Ala Thr Gln Asn Ala Ser Val Tyr Leu Glu Asn Thr Trp Phe Trp
   740                745                750
Thr Ala Asp His Asp Leu Asp Ala Ser Gly His Gly Gln Ile Asn Ile
   755                760                765
Tyr Asn Gly Arg Gly Val Leu Leu Glu Ser Glu Gly Pro Ala Trp Leu
   770                775                780
Tyr Gly Thr Ala Ala Glu His Asn Gln Leu Tyr Asn Tyr Gln Val Ser
785                790                795                800
Asn Ala Lys Asn Val Phe Met Gly Leu Ile Gln Thr Glu Thr Pro Tyr
   805                810                815
Tyr Gln Ala Asn Pro Asn Ala Leu Thr Pro Phe Thr Pro Gln Thr Ser
   820                825                830
Trp Asn Asp Pro Asp Phe Ser Tyr Cys Thr Thr Asp Ser Cys Arg Lys
   835                840                845
Ala Trp Gly Leu Arg Val Ile Asn Ser Ser Glu Thr Tyr Ile Tyr Gly
   850                855                860
Ala Gly Leu Tyr Ser Phe Phe Asp Asn Tyr Ser Gln Thr Cys Leu Ala
865                870                875                880
Thr Glu Ser Cys Gln Glu Asn Ile Val Glu Val Asp Cys Ser Asp Val
   885                890                895
His Ile Tyr Gly Leu Ser Thr Lys Ala Ser Thr Asn Met Ile Thr Ser
   900                905                910
Ser Ser Gly Asp Gly Leu Val Pro Gln Glu Gly Asn Glu Ser Thr Tyr
   915                920                925
Cys Ser Thr Ile Ala Leu Phe Glu Gln Ser Ser Ser Ile
   930                935                940

```