(19) DANMARK

(10) DK/EP 3487996 T5

(12) Rettet oversættelse af europæisk patentskrift

Patent- og Varemærkestyrelsen

(51) Int.Cl.: *C 12 N 9/58 (2006.01)*

(45) Oversættelsen bekendtgjort den: **2024-09-02**

(80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2024-04-10**

(86) Europæisk ansøgning nr.: **17746005.2**

(86) Europæisk indleveringsdag: **2017-07-14**

(87) Den europæiske ansøgnings publiceringsdag: **2019-05-29**

(86) International ansøgning nr.: **EP2017067883**

(87) Internationalt publikationsnr.: **WO2018015304**

(30) Prioritet: **2016-07-21 EP 16180497** **2016-10-21 EP 16195078**

(84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**

(73) Patenthaver: **Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Danmark**

(72) Opfinder: **AYABE, Keiichi, Makuhari Techno Garden CB-6, 3 Nakase 1-chome, Mihamaku-ku, Chiba-shi 261-8501, Japan**
MATSUI, Tomoko, Makuhari Techno Garden CB-6, 3 Nakase 1-chome, Mihamaku-ku, Chiba-shi 261-8501, Japan
TOMIKI,Aki, Makuhari Techno Garden CB-6, 3 Nakase 1-chome, Mihamaku-ku, Chiba-shi 261-8501, Japan
KURAKATA, Yuma, Makuhari Techno Garden CB-6, 3 Nakase 1-chome, Mihamaku-ku, Chiba-shl 261-8501, Japan
FRIIS, Esben, P., Krogshoejvej 36, 2880 Bagsværd, Danmark
NIELSEN, Jens, E., Krogshoejvej 36, 2880 Bagsværd, Danmark
PACHE, Roland, Alexander, Krogshoejvej 36, 2880 Bagsværd, Danmark

(54) Benævnelse: **Serinproteasevarianter og polynukleotider, der koder for samme**

(56) Fremdragne publikationer:
EP-B1- 3 487 995
WO-A1-2014/037438
WO-A1-2017/050291

DESCRIPTION

Description

REFERENCE TO A SEQUENCE LISTING

[0001] This application contains a Sequence Listing in computer readable form.

BACKGROUND OF THE INVENTION

[0002] Production of fermentation products, such as ethanol, from starch-containing material is well-known in the art. Generally, two different kinds of processes are used. The most commonly used process, often referred to as a "conventional process", includes liquefying gelatinized starch at high temperature using typically a bacterial alpha-amylase, followed by simultaneous saccharification and fermentation carried out in the presence of a glucoamylase and a fermenting organism. Another well-known process, often referred to as a "raw starch hydrolysis"-process (RSH process) includes simultaneously saccharifying and fermenting granular starch below the initial gelatinization temperature typically in the presence of an acid fungal alpha-amylase and a glucoamylase.

[0003] US Patent No. 5,231,017-A discloses the use of an acid fungal protease during ethanol fermentation in a process comprising liquefying gelatinized starch with an alpha-amylase.

[0004] WO 2003/066826 discloses a raw starch hydrolysis process (RSH process) carried out on non-cooked mash in the presence of fungal glucoamylase, alpha-amylase and fungal protease.

[0005] WO 2007/145912 discloses a process for producing ethanol comprising contacting a slurry comprising granular starch obtained from plant material with an alpha-amylase capable of solubilizing granular starch at a pH of 3.5 to 7.0 and at a temperature below the starch gelatinization temperature for a period of 5 minutes to 24 hours; obtaining a substrate comprising greater than 20% glucose, and fermenting the substrate in the presence of a fermenting organism and starch hydrolyzing enzymes at a temperature between 10°C and 40°C for a period of 10 hours to 250 hours. Additional enzymes added during the contacting step may include protease.

[0006] WO 2014/037438 discloses serine proteases derived from *Meripilus giganteus, Trametes versicolor, and Dichomitus squalens* and their use in animal feed.

[0007] US provisional application 62/232,903 discloses the use of the *Meripilus giganteus* S53 protease in the saccharification and/or fermentation step in a starch to ethanol process.

[0008] It is an object of the present invention to identify variants of the *M. giganteus* S53 proteases that will result in increased storage stability, in particular an increased thermo-stability of the variant protease compared to the wild type parent enzyme.

[0009] The present disclosure provides protease variants with improved properties compared to its parent.

SUMMARY OF THE INVENTION

[0010] In a first aspect the present invention relates to a protease variant comprising a modification at one or more position corresponding to positions 39, 50, 57, 60, 74, 81, 84, 109, 110, 111, 115, 117, 124, 142, 145, 146, 154, 182, 183, 187, 207, 209, 210, 228, 267, 271, 272, 274, 280, 294, 317, 318, 320, 321, 322, 328, 343, 348, 362 or 363 of the polypeptide of SEQ ID NO: 3, wherein the variant has protease activity and wherein the variant has at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%, but less than 100% sequence identity to the mature polypeptide of SEQ ID NO: 3, and wherein the variant has increased thermo-stability compared to the protease of SEQ ID NO: 3, wherein the increased thermo-stability is increased residual activity measured after incubation for 30 min at a temperature in the range from 55 to 60 degrees Celsius, and wherein the variant comprises at least one of the following modifications or combination of modifications:

N115L;

S183P;

N115L + Q182G;

N115L + Q182R;

E81R + S183P;

E81K + S183P;

S183P + Q154V;

S183P + Q142W;

Q142R + S183P;

S183P + T146A;

S183P + T146W;

S183P + I228R;

S183P + D267N;

S183P + S272V;

S183P + S272R;

T146Y + S183P;

S183P + S294A;

S183P + T362A;

S183P + S294A;

S183P + E74W;

S183P + E81A;

N115L + S183L + S187L;

S183L + V209L + S210deletion;

D109P + V209L + S210 deletion;

N115D + V209L + S210 deletion;

E81R + V209L + S210 deletion;

D109P + V209L + S210 deletion;

N115D + V209L + S210 deletion;

E81R + V209L + S210 deletion;

T146W + S183P + D280N;

I84C + S183P + S272C;

I39M + Q142R + S183P;

I39R + Q142R + S183P;

I39L + Q142R + S183P;

I39C + Q142R + S183P;

E117D + Q142R + S183P;

S60D + Q142R + S183P;

N115L + S183L +S187L + P348A;

D109P + S183P + V209L + S210 deletion;

N115D + S183P + V209L + S210 deletion;

E81R + S183P + V209L +S210 deletion;

V209L + S210 deletion + S317A + S318 deletion;

Q142R + N145G + T146E + S183P;

Q142R + N145Q + T146D + S183P;

Q142R + N145V + T146E + S183P;

Q142R + N145D + T146E + S183P;

Q142R + N145K + T146E + S183P;

Q142R + N145A + T146D +S183P;

Q142R + N145E + T146E + S183P;

N115L + S183L + S187L + V209W + S210 deletion;

N115L + S183L + S187L + V209L + S210 deletion;

N115L + S183L + S187L + S317G + S318 deletion;

N115L + S183L + S187L + S317S + S318 deletion;

N115L +S183L + S187L + S317A + S318 deletion;

E81R + V209L + S210 deletion + S317A + S318 deletion.

[0011] The present invention also relates to polynucleotides encoding the variants; nucleic acid constructs, vectors, and host cells comprising the polynucleotides; and methods of producing the variants. In a further aspect the present invention relates to compositions comprising the variants of the invention.

[0012] The present invention also relates to a process for producing a fermentation product from starch-containing material comprising simultaneously saccharifying and fermenting starch-containing material using a carbohydrate-source generating enzymes and a fermenting organism at a temperature below the initial gelatinization temperature of said starch-containing material in the presence of a variant protease. In another aspect the present invention relates to a process for producing a fermentation product from starch-containing material comprising the steps of: (a) liquefying starch-containing material in the presence of an alpha-amylase; (b) saccharifying the liquefied material obtained in step (a) using a glucoamylse; (c)

fermenting using a fermenting organism; wherein a variant protease of the invention is present during step b) and/or c).

DEFINITIONS

[0013] Protease: The term "protease" (also designated peptidases, proteinases, peptide hydrolases, or proteolytic enzymes) means a proteolytic activity (EC 3.4) that catalyzes the cleavage of peptide bonds. For purposes of the present invention, serine protease activity is determined according to the procedure described in the Examples. In one aspect, the variants of the present invention have at least 20%, e.g., at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the protease activity of the mature polypeptide of SEQ ID NO: 2.

[0014] Protease activity: The term "protease activity" means proteolytic activity (EC 3.4). There are several protease activity types such as trypsin-like proteases cleaving at the carboxyterminal side of Arg and Lys residues and chymotrypsin-like proteases cleaving at the carboxyterminal side of hydrophobic amino acid residues. Proteases of the invention are serine endopeptidases (EC 3.4.21) with acidic pH-optimum (pH optimum < pH 7).

[0015] Protease activity can be measured using any assay, in which a substrate is employed, that includes peptide bonds relevant for the specificity of the protease in question. Assay-pH and assay-temperature are likewise to be adapted to the protease in question. Examples of assay-pH-values are pH 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12. Examples of assay-temperatures are 15, 20, 25, 30, 35, 37, 40, 45, 50, 55, 60, 65, 70, 80, 90, or 95°C. Examples of general protease substrates are casein, bovine serum albumin and haemoglobin. In the classical Anson and Mirsky method, denatured haemoglobin is used as substrate and after the assay incubation with the protease in question, the amount of trichloroacetic acid soluble haemoglobin is determined as a measurement of protease activity (Anson, M.L. and Mirsky, A.E., 1932, J. Gen. Physiol. 16: 59 and Anson, M.L., 1938, J. Gen. Physiol. 22: 79).

[0016] For the purpose of the present invention, protease activity was determined using assays which are described in "Materials and Methods", such as the Kinetic Suc-AAPF-pNA assay, Protazyme AK assay, Kinetic Suc-AAPX-pNA assay and o-Phthaldialdehyde (OPA). For the Protazyme AK assay, insoluble Protazyme AK (Azurine-Crosslinked Casein) substrate liberates a blue colour when incubated with the protease and the colour is determined as a measurement of protease activity. For the Suc-AAPF-pNA assay, the colourless Suc-AAPF-pNA substrate liberates yellow paranitroaniline when incubated with the protease and the yellow colour is determined as a measurement of protease activity.

[0017] Endo-protease/Exo-proteases: Polypeptides having protease activity, or proteases, are sometimes also designated peptidases, proteinases, peptide hydrolases, or proteolytic enzymes. Proteases may be of the exo-type (exopeptidases) that hydrolyse peptides starting at either end thereof, or of the endo-type that act internally in polypeptide chains

(endopeptidases).

[0018] S53 protease: The term "S53 " means a protease activity selected from:

1. (a) proteases belonging to the EC 3.4.21 enzyme group; and/or
2. (b) proteases belonging to the EC 3.4.14 enzyme group; and/or
3. (c) Serine proteases of the peptidase family S53 that comprises two different types of peptidases: tripeptidyl aminopeptidases (exo-type) and endo-peptidases; as described in 1993, Biochem. J. 290:205-218 and in MEROPS protease database, release, 9.4 (31 January 2011) (www.merops.ac.uk). The database is described in Rawlings, N.D., Barrett, A.J. and Bateman, A., 2010, "MEROPS: the peptidase database", Nucl. Acids Res. 38: D227-D233.

[0019] For determining whether a given protease is a Serine protease, and a family S53 protease, reference is made to the above Handbook and the principles indicated therein. Such determination can be carried out for all types of proteases, be it naturally occurring or wild-type proteases; or genetically engineered or synthetic proteases.

[0020] Allelic variant: The term "allelic variant" means any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequences. An allelic variant of a polypeptide is a polypeptide encoded by an allelic variant of a gene.

[0021] cDNA: The term "cDNA" means a DNA molecule that can be prepared by reverse transcription from a mature, spliced, mRNA molecule obtained from a eukaryotic or prokaryotic cell. cDNA lacks intron sequences that may be present in the corresponding genomic DNA. The initial, primary RNA transcript is a precursor to mRNA that is processed through a series of steps, including splicing, before appearing as mature spliced mRNA.

[0022] Coding sequence: The term "coding sequence" means a polynucleotide, which directly specifies the amino acid sequence of a variant. The boundaries of the coding sequence are generally determined by an open reading frame, which begins with a start codon such as ATG, GTG or TTG and ends with a stop codon such as TAA, TAG, or TGA. The coding sequence may be a genomic DNA, cDNA, synthetic DNA, or a combination thereof.

[0023] Control sequences: The term "control sequences" means nucleic acid sequences necessary for expression of a polynucleotide encoding a variant of the present invention. Each control sequence may be native (*i.e.*, from the same gene) or foreign (*i.e.*, from a different gene) to the polynucleotide encoding the variant or native or foreign to each other. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a

minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the polynucleotide encoding a variant.

[0024] Expression: The term "expression" includes any step involved in the production of a variant including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

[0025] Expression vector: The term "expression vector" means a linear or circular DNA molecule that comprises a polynucleotide encoding a variant and is operably linked to control sequences that provide for its expression.

[0026] Fragment: The term "fragment" means a polypeptide having one or more (e.g., several) amino acids absent from the amino and/or carboxyl terminus of a mature polypeptide; wherein the fragment has protease activity.

[0027] High stringency conditions: The term "high stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 50% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 65°C.

[0028] Host cell: The term "host cell" means any cell type that is susceptible to transformation, transfection, transduction, or the like with a nucleic acid construct or expression vector comprising a polynucleotide of the present invention. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

[0029] Improved property: The term "improved property" means a characteristic associated with a variant that is improved compared to the parent. Such improved properties include, but are not limited to, increased stability under storage conditions, increased thermo-stability, and increased residual activity.

[0030] Isolated: The term "isolated" means a substance in a form or environment which does not occur in nature. Non-limiting examples of isolated substances include (1) any non-naturally occurring substance, (2) any substance including, but not limited to, any enzyme, variant, nucleic acid, protein, peptide or cofactor, that is at least partially removed from one or more or all of the naturally occurring constituents with which it is associated in nature; (3) any substance modified by the hand of man relative to that substance found in nature; or (4) any substance modified by increasing the amount of the substance relative to other components with which it is naturally associated (*e.g.*, multiple copies of a gene encoding the substance; use of a stronger promoter than the promoter naturally associated with the gene encoding the substance). An isolated substance may be present in a fermentation broth sample.

[0031] Low stringency conditions: The term "low stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 25% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 50°C.

[0032] Mature polypeptide: The term "mature polypeptide" means a polypeptide in its final form following translation and any post-translational modifications, such as N-terminal processing, C-terminal truncation, glycosylation, phosphorylation, etc. In one aspect, the mature polypeptide is amino acids 199 to 564 of SEQ ID NO: 2. Amino acids 1 to 17 of SEQ ID NO: 2 are a signal peptide, and amino acids 18 to 198 are a propeptide. The N-terminals of the mature S53 polypeptides used according to the present invention were experimentally confirmed based on EDMAN N-terminal sequencing data and Intact MS data. The mature polypeptides are also included as SEQ ID NO: 3 (mature S53 protease 3 from *Meripilus giganteus.* It is known in the art that a host cell may produce a mixture of two of more different mature polypeptides (*i.e.*, with a different C-terminal and/or N-terminal amino acid) expressed by the same polynucleotide.

[0033] Mature polypeptide coding sequence: The term "mature polypeptide coding sequence" means a polynucleotide that encodes a mature polypeptide having protease activity. In one aspect, the mature polypeptide coding sequence is nucleotides 595 to 1692 of SEQ ID NO: 1. Nucleotides 1 to 51 of SEQ ID NO: 1 encode a signal peptide, nucleotides 52 to 594 encode a propeptide.

[0034] Medium stringency conditions: The term "medium stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 35% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 55°C.

[0035] Medium-high stringency conditions: The term "medium-high stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 35% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 60°C.

[0036] Modification: The term "modification(s)" is in the context of the present invention to be understood as a substitution, insertion, and/or deletion, at one or more (e.g., several) positions. A substitution means replacement of the amino acid occupying a position with a different amino acid; a deletion means removal of the amino acid occupying a position; and an insertion means adding an amino acid adjacent to and immediately following the amino acid occupying a position.

[0037] Mutant: The term "mutant" means a polynucleotide encoding a variant.

[0038] Nucleic acid construct: The term "nucleic acid construct" means a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature or which is synthetic, which comprises one or more control sequences. In one embodiment the one or more control sequences are heterologous (of different origin/species) to the coding sequence encoding the polypeptide of the invention.

[0039] Operably linked: The term "operably linked" means a configuration in which a control sequence is placed at an appropriate position relative to the coding sequence of a polynucleotide such that the control sequence directs expression of the coding sequence.

[0040] Parent or parent protease: The term "parent" or "parent protease" means any polypeptide with protease activity to which an alteration is made to produce the enzyme variants of the present invention.

[0041] Sequence identity: The relatedness between two amino acid sequences or between two nucleotide sequences is described by the parameter "sequence identity".

[0042] For purposes of the present invention, the sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends Genet. 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

(Identical Residues × 100)/(Length of Alignment - Total Number of Gaps in Alignment)

[0043] For purposes of the present invention, the sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *supra)* as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.,* 2000, *supra),* preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

(Identical Deoxyribonucleotides x 100)/(Length of Alignment – Total Number of Gaps in Alignment)

[0044] Subsequence: The term "subsequence" means a polynucleotide having one or more (*e.g.*, several) nucleotides absent from the 5' and/or 3' end of a mature polypeptide coding sequence; wherein the subsequence encodes a fragment having protease activity. In one aspect, a subsequence contains at least 1098 nucleotides (*e.g.,* nucleotides 595 to 1692 of SEQ ID NO: 1).

[0045] Variant: The term "variant" means a polypeptide having protease activity comprising a modification(s) at one or more (*e.g.,* several) positions. The variants of the present invention have at least 20%, *e.g.,* at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the protease activity of the mature polypeptide of SEQ ID NO: 2, disclosed herein as SEQ ID NO: 3.

[0046] Very high stringency conditions: The term "very high stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 50% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 70°C.

[0047] Very low stringency conditions: The term "very low stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 25% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 45°C.

[0048] Wild-type protease: The term "wild-type" protease means a protease expressed by a naturally occurring microorganism, such as a bacterium, yeast, or filamentous fungus found in nature.

Conventions for Designation of Variants

[0049] For purposes of the present invention, the mature polypeptide comprised in SEQ ID NO: 2 is used to determine the corresponding amino acid residue in another protease. The amino acid sequence of another protease is aligned with the mature polypeptide comprised in SEQ ID NO: 2 (disclosed herein as SEQ ID NO: 3), and based on the alignment, the amino acid position number corresponding to any amino acid residue in the mature polypeptide comprised in SEQ ID NO: 2 is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends Genet. 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix.

[0050] Identification of the corresponding amino acid residue in another protease than the *Meripilus giganteus* S53 protease can be determined by an alignment of multiple polypeptide sequences using several computer programs including, but not limited to, MUSCLE (multiple sequence comparison by log-expectation; version 3.5 or later; Edgar, 2004, Nucleic Acids Research 32: 1792-1797), MAFFT (version 6.857 or later; Katoh and Kuma, 2002, Nucleic Acids Research 30: 3059-3066; Katoh et al., 2005, Nucleic Acids Research 33: 511-518; Katoh and Toh, 2007, Bioinformatics 23: 372-374; Katoh et al., 2009, Methods in Molecular Biology 537: 39-64; Katoh and Toh, 2010, Bioinformatics 26: 1899-1900), and EMBOSS EMMA employing ClustalW (1.83 or later; Thompson et al., 1994, Nucleic Acids Research 22: 4673-4680), using their respective default parameters.

[0051] When the other enzyme has diverged from the mature polypeptide of SEQ ID NO: 2 such that traditional sequence-based comparison fails to detect their relationship (Lindahl and Elofsson, 2000, J. Mol. Biol. 295: 613-615), other pairwise sequence comparison algorithms can be used. Greater sensitivity in sequence-based searching can be attained using search programs that utilize probabilistic representations of polypeptide families (profiles) to search databases. For example, the PSI-BLAST program generates profiles through an iterative database search process and is capable of detecting remote homologs (Atschul et al., 1997, Nucleic Acids Res. 25: 3389-3402). Even greater sensitivity can be achieved if the family or superfamily for the polypeptide has one or more representatives in the protein structure databases. Programs such as GenTHREADER (Jones, 1999, J. Mol. Biol. 287: 797-815; McGuffin and Jones, 2003, Bioinformatics 19: 874-881) utilize information from a variety of sources (PSI-BLAST, secondary structure prediction, structural alignment profiles, and solvation potentials) as input to a neural network that predicts the structural fold for a query sequence. Similarly, the method of Gough et al., 2000, J. Mol. Biol. 313: 903-919, can be used to align a sequence of unknown structure with the superfamily models present in the SCOP database. These alignments can in turn be used to generate homology models for the polypeptide, and such models can be assessed for accuracy using a variety of tools developed for that purpose.

[0052] For proteins of known structure, several tools and resources are available for retrieving and generating structural alignments. For example the SCOP superfamilies of proteins have been structurally aligned, and those alignments are accessible and downloadable. Two or more protein structures can be aligned using a variety of algorithms such as the distance alignment matrix (Holm and Sander, 1998, Proteins 33: 88-96) or combinatorial extension (Shindyalov and Bourne, 1998, Protein Engineering 11: 739-747), and implementation of these algorithms can additionally be utilized to query structure databases with a structure of interest in order to discover possible structural homologs (*e.g.*, Holm and Park, 2000, Bioinformatics 16: 566-567).

[0053] In describing the variants of the present invention, the nomenclature described below is adapted for ease of reference. The accepted IUPAC single letter or three letter amino acid abbreviation is employed.

[0054] Substitutions. For an amino acid substitution, the following nomenclature is used: Original amino acid, position, substituted amino acid. Accordingly, the substitution of threonine at position 226 with alanine is designated as "Thr226Ala" or "T226A". Multiple mutations are separated by addition marks ("+"), *e.g.*, "Gly205Arg + Ser411Phe" or "G205R + S411F", representing substitutions at positions 205 and 411 of glycine (G) with arginine (R) and serine (S) with phenylalanine (F), respectively.

[0055] Deletions. For an amino acid deletion, the following nomenclature is used: Original amino acid, position, *. Accordingly, the deletion of glycine at position 195 is designated as "Gly195*" or "G195*". Multiple deletions are separated by addition marks ("+"), *e.g.*, "Gly195* + Ser411*" or "G195* + S411*".

[0056] Insertions. For an amino acid insertion, the following nomenclature is used: Original amino acid, position, original amino acid, inserted amino acid. Accordingly the insertion of lysine after glycine at position 195 is designated "Gly195GlyLys" or "G195GK". An insertion of multiple amino acids is designated [Original amino acid, position, original amino acid, inserted amino acid #1, inserted amino acid #2; etc.]. For example, the insertion of lysine and alanine after glycine at position 195 is indicated as "Gly195GlyLysAla" or "G195GKA".

[0057] In such cases the inserted amino acid residue(s) are numbered by the addition of lower case letters to the position number of the amino acid residue preceding the inserted amino acid residue(s). In the above example, the sequence would thus be:

Parent:	Variant:
195	195 195a 195b
G	G - K - A

[0058] Multiple modification. Variants comprising multiple modifications are separated by addition marks ("+"), e.g., "Arg170Tyr+Gly195Glu" or "R170Y+G195E" representing a substitution of arginine and glycine at positions 170 and 195 with tyrosine and glutamic acid, respectively.

[0059] Different modifications. Where different modification can be introduced at a position, the different modifications are separated by a comma, e.g., "Arg170Tyr,Glu" represents a substitution of arginine at position 170 with tyrosine or glutamic acid. Thus, "Tyr167Gly,Ala + Arg170Gly,Ala" designates the following variants:
"Tyr167Gly+Arg170Gly", "Tyr167Gly+Arg170Ala", "Tyr167Ala+Arg170Gly", and "Tyr167 Ala+Arg 170Ala".

DETAILED DESCRIPTION OF THE INVENTION

[0060] The present disclosure relates to protease variants, comprising a modification(s) at one

or more (*e.g.*, several) positions corresponding to specific positions of the mature polypeptide disclosed as SEQ ID NO: 3 (a parent protease), wherein the variant has protease activity. As explained herein the specific position numbers may change in case the mature parent protease is different from SEQ ID NO: 3. The improved properties of the variants of the invention falls in the following category, increased thermostability (measured as increase in thermal denaturation temperature, Td, and/or increased residual activity by the Suc-AAPF assay after incubation for 30 min at an elevated temperature in the range from 55 to 60 degrees Celsius as described in detail in the examples).

Variants

[0061] The present invention provides a protease variant comprising a modification at one or more position corresponding to positions 39, 50, 57, 60, 74, 81, 84, 109, 110, 111, 115, 117, 124, 142, 145, 146, 154, 182, 183, 187, 207, 209, 210, 228, 267, 271, 272, 274, 280, 294, 317, 318, 320, 321, 322, 328, 343, 348, 362 or 363 of the polypeptide of SEQ ID NO: 3, wherein the variant has protease activity and wherein the variant has at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%, but less than 100% sequence identity to the mature polypeptide of SEQ ID NO: 3, and wherein the variant has increased thermo-stability compared to the protease of SEQ ID NO: 3, wherein the increased thermo-stability is increased residual activity measured after incubation for 30 min at an elevated temperature in the range from 55 to 60 degrees Celsius, and wherein the variant comprises at least one of the following modifications or combination of modifications:

N115L;

S183P;

N115L + Q182G;

N115L + Q182R;

E81R + S183P;

E81K + S183P;

S183P + Q154V;

S183P + Q142W;

Q142R + S183P;

S183P + T146A;

S183P + T146W;

S183P + I228R;

S183P + D267N;

S183P + S272V;

S183P + S272R;

T146Y + S183P;

S183P + S294A;

S183P + T362A;

S183P + S294A;

S183P + E74W;

S183P + E81A;

N115L + S183L + S187L;

S183L + V209L + S210 deletion;

D109P + V209L + S210 deletion

N115D + V209L + S210 deletion;

E81R + V209L + S210 deletion;

D109P + V209L + S210 deletion;

N115D + V209L + S210 deletion;

E81R + V209L + S210 deletion;

T146W + S183P + D280N;

I84C + S183P + S272C;

I39M + Q142R + S183P;

I39R + Q142R + S183P;

I39L + Q142R + S183P;

I39C + Q142R + S183P;

E117D + Q142R + S183P;

S60D + Q142R + S183P;

N115L + S183L +S187L + P348A;

D109P + S183P + V209L + S210 deletion;

N115D + S183P + V209L + S210 deletion;

E81R + S183P + V209L +S210 deletion;

V209L + S210 deletion + S317A + S318 deletion;

Q142R + N145G + T146E + S183P;

Q142R + N145Q + T146D + S183P;

Q142R + N145V + T146E + S183P ;

Q142R + N145D + T146E + S183P;

Q142R + N145K + T146E + S183P;

Q142R + N145A + T146D +S183P;

Q142R + N145E + T146E + S183P;

N115L + S183L + S187L + V209W+ S210 deletion;

N115L + S183L + S187L + V209L + S210 deletion;

N115L + S183L + S187L + S317G + S318 deletion;

N115L + S183L + S187L + S317S + S318 deletion;

N115L +S183L + S187L + S317A + S318 deletion;

E81R + V209L + S210 deletion + S317A + S318 deletion.

More particularly the variants have a residual activity of at least 10%, particularly at least 12%, more particularly at least 15%, measured after incubation for 30 minutes at 56°C.

[0062] In a further specific embodiment the variant comprises at least one of the following modifications or combination of modifications:

N115L + Q182G;

Q142R +S183P;

Q142R +N145G +T146E +S183P;

Q142R +N145Q + T146D + S183P;

Q142R +N145V + T146E +S183P;

Q142R +N145D +T146E + S183P;

Q142R +N145K + T146E + S183P;

Q142R +N145A + T146D +S183P;

I39M + Q142R +S183P;

Q142R + N145E + T146E +S183P;

I39R + Q142R + S183P;

I39L + Q142R + S183P;

E117D + Q142R + S183P;

S60D + Q142R + S183P; and wherein the variant has residual activity of at least 30% measured after incubation for 30 min at an elevated temperature of 57 degrees Celsius.

[0063] In a further specific embodiment the variant comprises at least one of the following modifications or combination of modifications:

Q142R+ S183P;

Q142R +N145G +T146E +S183P;

Q142R +N145Q +T146D +S183P;

Q142R + N145V + T146E +S183P;

Q142R + N145D + T146E + S183P;

Q142R + N145K + T146E + S183P;

Q142R + N145A + T146D +S183P;

I39M + Q142R + S183P;

Q142R + N145E + T146E + S183P;

I39R + Q142R + S183P;

I39L +Q142R +S183P;

E117D +Q142R S183P;

Q142R + S183P;

S60D + Q142R + S183P;

Q142R +S183P; and wherein the variant has residual activity of at least 70% measured after incubation for 30 min at 57 degrees Celsius.

[0064] In a further specific embodiment the variant comprises at least one of the following modification or combination of modifications:

Q142R + S183P;

I39C + Q142R + S183P;

E117D + Q142R + S183P;

Q142R + S183P;

S60D + Q142R + S183P; and wherein the variant has residual activity of at least 40% measured after incubation for 30 min at 60 degrees Celsius.

[0065] In a further specific embodiment the variant comprises at least one of the following modification or combination of modifications:

Q142R + S183P;

I39C + Q142R +S183P; and wherein the variant has residual activity of at least 40% measured after incubation for 30 minutes at 62°C.

[0066] The present invention relates to a variant comprising a modification at position corresponding to position 50, 57, 81, 84, 109, 110, 111, 124, 142, 145, 146, 154, 182, 183, 207, 209, 210, 228, 267, 271, 272, 274, 280, 294, 317, 318, 320, 321, 322, 328, 343, 362, or 363 of the polypeptide of SEQ ID NO: 3, wherein the variant has protease activity and wherein the variant has at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%, but less than 100% sequence identity to the polypeptide of SEQ ID NO: 3, wherein the increase in thermo-stability is an increase in thermal denaturation temperature measured by TSA. In particular, the increased thermo-stability measured as Td by TSA assay is at least 59°C, the variant comprises at least one of the following modifications or combination of modifications:

S183P;

K57R +S183P;

D109P +S183P+ V209L +S210 deletion;

E81R+ S183P +V209L +S210 deletion;

E81R +V209L +S210 deletion;

Q154V +S183P;

Q142W +S183P;

Q142R +S183P;

T146A +S183P;

T146W +S183P;

S183P +I228R;

S183P +D267N;

S183P +S272V;

E81R +V209L +S210 deletion +S317A +S318 deletion;

S183P +T328C +K343C;

S183P +G320C+ A363C;

T146W +S183L D+280N;

T146W +S183P +D280N;

T146Y +S183P;

S183P +Q207R;

S50C +S183P +V271C;

I84C +S183P+ S272C;

Q142W +T146W +S183P;

Q142W+T146W+S183P+ D280N;

S183P +S294A;

S183P +K321G;

S183P +T362A;

Q182G;

Q142W +T146W +Q182R;

S272V;

D109N +D110N;

F111P;

S318N +K321A +A322S;

E81R+ T146W;

E81R +Q142R +S183P;

E81R +Q142W+S183P

S183P +G274G;

E81R;

N124L +Q142R +S183P;

N124W +Q142R +S183P;

N124Q +Q142R +S183P;

Q142R + N145V + T146E +S183P;

Q142R +N145D +T146E + S183P;

Q142R + N145A +T146D +S183P; and wherein the increased thermo-stability measured as Td by TSA assay is at least 59°C.

[0067] In a further specific embodiment, the variant comprises at least one of the following modifications or combination of modifications:

D109P + S183P + V209L + S210 deletion;

E81R + S183P + V209L + S210 deletion;

E81R + V209L +S210 deletion;

Q154V + S183P;

Q142W +S183P;

Q142R + S183P;

T146A + S183P;

T146W + S183P;

S183P + D267N;

S183P + S272V;

E81R + V209L + S210 deletion+ S317A +S318 deletion;

T146W + S183L +D280N;

T146W + S183P +D280N;

T146Y + S183P;

S183P + Q207R;

S50C + S183P + V271C;

I84C + S183P + S272C;

Q142W + T146W + S183P;

Q142W + T146W + S183P + D280N;

S183P + S294A;

Q142W + T146W + Q182R;

S272V;

E81R + T146W;

E81R + Q142R + S183P;

E81R + Q142W + S183P;

S183P + G274G;

E81R;

N124L + Q142R + S183P;

N124W + Q142R + S183P;

N124Q + Q142R + S183P;

Q142R + N145V + T146E + S183P;

Q142R + N145D + T146E +S183P;

Q142R + N145A + T146D + S183P; and wherein the increased thermo-stability measured as Td by TSA assay is at least 61°C.

[0068] In a further specific embodiment, the variant comprises at least one of the following modifications or combination of modifications:

E81R + S183P+ V209L +S210 deletion;

Q142R + S183P;

T146W+S183L + D280N;

T146W + S183P+ D280N;

S50C + S183P + V271C;

I84C + S183P + S272C;

Q142W + T146W + S183P+ D280N;

S272V;

E81R +T146W;

E81R + Q142R + S183P;

N124L + Q142R +S183P;

N124W + Q142R +S183P;

N124Q + Q142R + S183P;

Q142R + N145V + T146E + S183P;

Q142R + N145D + T146E + S183P;

Q142R + N145A + T146D + S183P; and wherein the increased thermo-stability measured as Td by TSA assay is at least 63°C.

[0069] In a further specific embodiment, the variant comprises at least one of the following modifications or combination of modifications:

Q142R + S183P;

S50C + S183P + V271C;

E81R + Q142R + S183P;

N124L + Q142R + S183P;

N124Q + Q142R + S183P;

Q142R + N145V + T146E + S183P;

Q142R + N145D + T146E + S183P; or

Q142R + N145A + T146D + S183P; and wherein the increased thermo-stability measured as Td by TSA assay is at least 65°C.

[0070] The variants may further comprise one or more additional modification(s) at one or more (*e.g.*, several) other positions. Such further modifications may preferably not change the properties of the protease variants of the present invention.

[0071] The amino acid changes may be of a minor nature, that is conservative amino acid substitutions or insertions that do not significantly affect the folding and/or activity of the protein; small deletions, typically of 1-30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to 20-25 residues; or a small extension that facilitates purification by changing net charge or another function, such as a poly-histidine tract, an antigenic epitope or a binding domain.

[0072] Therefore even though the protease variants according to the invention may only comprise one specific substitution providing the improved property according to the invention it may still have addition modifications leading to a variant protease having at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%, but less than 100% sequence identity, to the amino acid sequence of the mature parent protease, e.g., the protease of SEQ ID NO: 3. These additional modification should preferably not significantly change the improved properties of the variant protease.

[0073] Examples of conservative substitutions are within the groups of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine and asparagine), hydrophobic amino acids (leucine, isoleucine and valine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine, serine, threonine and methionine). Amino acid substitutions that do not generally alter specific activity are known in the art and are described, for example, by H. Neurath and R.L. Hill, 1979, In, The Proteins, Academic Press, New York. Common substitutions are Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, and Asp/Gly.

[0074] Essential amino acids in a polypeptide can be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, 1989, Science 244: 1081-1085). In the latter technique, single alanine mutations are introduced at every residue in the molecule, and the resultant mutant molecules are tested for protease activity to identify amino acid residues that are critical to the activity of the molecule. See also, Hilton et al., 1996, J. Biol. Chem. 271: 4699-4708. The active site of the enzyme or other biological interaction can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic resonance, crystallography, electron diffraction, or photoaffinity labeling, in conjunction with mutation of putative contact site amino acids. See, for example, de Vos et al., 1992, Science 255: 306-312; Smith et al., 1992, J. Mol. Biol. 224: 899-904; Wlodaver et al., 1992, FEBS Lett. 309: 59-64.

[0075] In an embodiment, the variant has improved (increased) thermo-stability compared to the parent enzyme, e.g., the polypeptide of SEQ ID NO: 3.

[0076] In an embodiment, the variant has improved (increased) residual activity compared to parent enzyme, e.g., the polypeptide of SEQ ID NO: 3.

[0077] In an embodiment, the variant has improved (increased) thermal melting temperature compared to parent enzyme, e.g., the polypeptide of SEQ ID NO: 3.

Parent proteases

[0078] The parent protease may be (a) a polypeptide having at least 80% sequence identity to the mature polypeptide of SEQ ID NO: 2; (b) a polypeptide encoded by a polynucleotide that hybridizes under high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, (ii) the cDNA sequence thereof, or (iii) the full-length complement of (i) or (ii); or (c) a polypeptide encoded by a polynucleotide having at least 80% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1.

[0079] In an aspect, the parent has a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, which have protease activity. In one aspect, the amino acid sequence of the parent differs by up to 10 amino acids, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10, from the mature polypeptide of SEQ ID NO: 2.

[0080] In another aspect, the parent comprises or consists of the amino acid sequence of SEQ ID NO: 3. In another aspect, the parent comprises or consists of amino acids 199 to 564 of SEQ ID NO: 2.

[0081] In another aspect, the parent is encoded by a polynucleotide that hybridizes under high stringency conditions, or very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, (ii) the cDNA sequence thereof, or (iii) the full-length complement of (i) or (ii) (Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, New York).

[0082] The polynucleotide of SEQ ID NO: 1 or a subsequence thereof, as well as the polypeptide of SEQ ID NO: 2 or a fragment thereof, may be used to design nucleic acid probes to identify and clone DNA encoding a parent from strains of different genera or species according to methods well known in the art. In particular, such probes can be used for hybridization with the genomic DNA or cDNA of a cell of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Such probes can be considerably shorter than the entire sequence, but should be at least 15, *e.g.,* at least 25, at least 35, or at least 70 nucleotides in length. Preferably, the nucleic acid probe is at least 100 nucleotides in length, *e.g.,* at least 200 nucleotides, at least 300 nucleotides, at

least 400 nucleotides, at least 500 nucleotides, at least 600 nucleotides, at least 700 nucleotides, at least 800 nucleotides, or at least 900 nucleotides in length. Both DNA and RNA probes can be used. The probes are typically labeled for detecting the corresponding gene (for example, with ^{32}P, ^{3}H, ^{35}S, biotin, or avidin). Such probes are encompassed by the present invention.

[0083] A genomic DNA or cDNA library prepared from such other strains may be screened for DNA that hybridizes with the probes described above and encodes a parent. Genomic or other DNA from such other strains may be separated by agarose or polyacrylamide gel electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to and immobilized on nitrocellulose or other suitable carrier material. In order to identify a clone or DNA that hybridizes with SEQ ID NO: 1 or a subsequence thereof, the carrier material is used in a Southern blot.

[0084] For purposes of the present invention, hybridization indicates that the polynucleotide hybridizes to a labeled nucleic acid probe corresponding to (i) SEQ ID NO: 1; (ii) the mature polypeptide coding sequence of SEQ ID NO: 1; (iii) the cDNA sequence thereof; (iv) the full-length complement thereof; or (v) a subsequence thereof; under high to very high stringency conditions. Molecules to which the nucleic acid probe hybridizes under these conditions can be detected using, for example, X-ray film or any other detection means known in the art.

[0085] In one aspect, the nucleic acid probe is the mature polypeptide coding sequence of SEQ ID NO: 1. In another aspect, the nucleic acid probe is nucleotides 595 to 1692 of SEQ ID NO: 1. In another aspect, the nucleic acid probe is a polynucleotide that encodes the polypeptide of SEQ ID NO: 2; the mature polypeptide thereof; or a fragment thereof. In another aspect, the nucleic acid probe is SEQ ID NO: 1 or the cDNA sequence thereof.

[0086] In another embodiment, the parent is encoded by a polynucleotide having a sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 of at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%.

[0087] The polypeptide may be a hybrid polypeptide in which a region of one polypeptide is fused at the N-terminus or the C-terminus of a region of another polypeptide.

[0088] The parent may be a fusion polypeptide or cleavable fusion polypeptide in which another polypeptide is fused at the N-terminus or the C-terminus of the polypeptide of the present invention. A fusion polypeptide is produced by fusing a polynucleotide encoding another polypeptide to a polynucleotide of the present invention. Techniques for producing fusion polypeptides are known in the art, and include ligating the coding sequences encoding the polypeptides so that they are in frame and that expression of the fusion polypeptide is under control of the same promoter(s) and terminator. Fusion polypeptides may also be constructed using intein technology in which fusion polypeptides are created post-translationally (Cooper et al., 1993, EMBO J. 12: 2575-2583; Dawson et al., 1994, Science

266: 776-779).

[0089] A fusion polypeptide can further comprise a cleavage site between the two polypeptides. Upon secretion of the fusion protein, the site is cleaved releasing the two polypeptides. Examples of cleavage sites include, but are not limited to, the sites disclosed in Martin et al., 2003, J. Ind. Microbiol. Biotechnol. 3: 568-576; Svetina et al., 2000, J. Biotechnol. 76: 245-251; Rasmussen-Wilson et al., 1997, Appl. Environ. Microbiol. 63: 3488-3493; Ward et al., 1995, Biotechnology 13: 498-503; and Contreras et al., 1991, Biotechnology 9: 378-381; Eaton et al., 1986, Biochemistry 25: 505-512; Collins-Racie et al., 1995, Biotechnology 13: 982-987; Carter et al., 1989, Proteins: Structure, Function, and Genetics 6: 240-248; and Stevens, 2003, Drug Discovery World 4: 35-48.

[0090] In another aspect, the parent is a Meripilus giganteus S53 protease, *e.g.,* the protease of SEQ ID NO: 2 or the mature polypeptide thereof, disclosed herein as SEQ ID NO: 3.

Preparation of Variants

[0091] The variants can be prepared using any mutagenesis procedure known in the art, such as site-directed mutagenesis, synthetic gene construction, semi-synthetic gene construction, random mutagenesis, shuffling, etc.

[0092] Site-directed mutagenesis is a technique in which one or more (*e.g.,* several) mutations are introduced at one or more defined sites in a polynucleotide encoding the parent.

[0093] Site-directed mutagenesis can be accomplished *in vitro* by PCR involving the use of oligonucleotide primers containing the desired mutation. Site-directed mutagenesis can also be performed *in vitro* by cassette mutagenesis involving the cleavage by a restriction enzyme at a site in the plasmid comprising a polynucleotide encoding the parent and subsequent ligation of an oligonucleotide containing the mutation in the polynucleotide. Usually the restriction enzyme that digests the plasmid and the oligonucleotide is the same, permitting sticky ends of the plasmid and the insert to ligate to one another. See, *e.g.,* Scherer and Davis, 1979, Proc. Natl. Acad. Sci. USA 76: 4949-4955; and Barton et al., 1990, Nucleic Acids Res. 18: 7349-4966.

[0094] Site-directed mutagenesis can also be accomplished *in vivo* by methods known in the art. See, *e.g.,* U.S. Patent Application Publication No. 2004/0171154; Storici et al., 2001, Nature Biotechnol. 19: 773-776; Kren et al., 1998, Nat. Med. 4: 285-290; and Calissano and Macino, 1996, Fungal Genet. Newslett. 43: 15-16.

[0095] Any site-directed mutagenesis procedure can be used in the present invention. There are many commercial kits available that can be used to prepare variants.

[0096] Synthetic gene construction entails *in vitro* synthesis of a designed polynucleotide molecule to encode a polypeptide of interest. Gene synthesis can be performed utilizing a

number of techniques, such as the multiplex microchip-based technology described by Tian et al. (2004, Nature 432: 1050-1054) and similar technologies wherein oligonucleotides are synthesized and assembled upon photo-programmable microfluidic chips.

[0097] Single or multiple amino acid substitutions, deletions, and/or insertions can be made and tested using known methods of mutagenesis, recombination, and/or shuffling, followed by a relevant screening procedure, such as those disclosed by Reidhaar-Olson and Sauer, 1988, Science 241: 53-57; Bowie and Sauer, 1989, Proc. Natl. Acad. Sci. USA 86: 2152-2156; WO 95/17413; or WO 95/22625. Other methods that can be used include error-prone PCR, phage display (e.g., Lowman et al., 1991, Biochemistry 30: 10832-10837; U.S. Patent No. 5,223,409; WO 92/06204) and region-directed mutagenesis (Derbyshire et al., 1986, Gene 46: 145; Ner et al., 1988, DNA 7: 127).

[0098] Mutagenesis/shuffling methods can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides expressed by host cells (Ness et al., 1999, Nature Biotechnology 17: 893-896). Mutagenized DNA molecules that encode active polypeptides can be recovered from the host cells and rapidly sequenced using standard methods in the art. These methods allow the rapid determination of the importance of individual amino acid residues in a polypeptide.

[0099] Semi-synthetic gene construction is accomplished by combining aspects of synthetic gene construction, and/or site-directed mutagenesis, and/or random mutagenesis, and/or shuffling. Semi-synthetic construction is typified by a process utilizing polynucleotide fragments that are synthesized, in combination with PCR techniques. Defined regions of genes may thus be synthesized *de novo,* while other regions may be amplified using site-specific mutagenic primers, while yet other regions may be subjected to error-prone PCR or non-error prone PCR amplification. Polynucleotide subsequences may then be shuffled.

Polynucleotides

[0100] The present invention also relates to polynucleotides encoding a variant of the present invention.

Nucleic Acid Constructs

[0101] The present invention also relates to nucleic acid constructs comprising a polynucleotide encoding a variant of the present invention operably linked to one or more control sequences that direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences.

[0102] The polynucleotide may be manipulated in a variety of ways to provide for expression of

a variant. Manipulation of the polynucleotide prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides utilizing recombinant DNA methods are well known in the art.

[0103] The control sequence may be a promoter, a polynucleotide which is recognized by a host cell for expression of the polynucleotide. The promoter contains transcriptional control sequences that mediate the expression of the variant. The promoter may be any polynucleotide that shows transcriptional activity in the host cell including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

[0104] Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a filamentous fungal host cell are promoters obtained from the genes for *Aspergillus nidulans* acetamidase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Aspergillus oryzae* TAKA amylase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Fusarium oxysporum* trypsin-like protease (WO 96/00787), *Fusarium venenatum* amyloglucosidase (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), *Rhizomucor miehei* lipase, *Rhizomucor miehei* aspartic proteinase, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase IV, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei* xylanase II, *Trichoderma reesei* beta-xylosidase, as well as the NA2-tpi promoter (a modified promoter from an *Aspergillus* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus* triose phosphate isomerase gene; non-limiting examples include modified promoters from an *Aspergillus* niger neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus nidulans* or *Aspergillus oryzae* triose phosphate isomerase gene); and mutant, truncated, and hybrid promoters thereof.

[0105] The control sequence may also be a transcription terminator, which is recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3'-terminus of the polynucleotide encoding the variant. Any terminator that is functional in the host cell may be used.

[0106] Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase, *Aspergillus oryzae* TAKA amylase, and *Fusarium oxysporum* trypsin-like protease.

[0107] The control sequence may also be an mRNA stabilizer region downstream of a promoter and upstream of the coding sequence of a gene which increases expression of the

gene.

[0108] The control sequence may also be a leader, a nontranslated region of an mRNA that is important for translation by the host cell. The leader sequence is operably linked to the 5'-terminus of the polynucleotide encoding the variant. Any leader that is functional in the host cell may be used.

[0109] Preferred leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase.

[0110] The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3'-terminus of the variant-encoding sequence and, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence that is functional in the host cell may be used.

[0111] Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase, *Aspergillus oryzae* TAKA amylase, and *Fusarium oxysporum* trypsin-like protease.

[0112] The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a variant and directs the variant into the cell's secretory pathway. The 5'-end of the coding sequence of the polynucleotide may inherently contain a signal peptide coding sequence naturally linked in translation reading frame with the segment of the coding sequence that encodes the variant. Alternatively, the 5'-end of the coding sequence may contain a signal peptide coding sequence that is foreign to the coding sequence. A foreign signal peptide coding sequence may be required where the coding sequence does not naturally contain a signal peptide coding sequence. Alternatively, a foreign signal peptide coding sequence may simply replace the natural signal peptide coding sequence in order to enhance secretion of the variant. However, any signal peptide coding sequence that directs the expressed variant into the secretory pathway of a host cell may be used.

[0113] Effective signal peptide coding sequences for filamentous fungal host cells are the signal peptide coding sequences obtained from the genes for *Aspergillus niger* neutral amylase, *Aspergillus* nigerglucoamylase, *Aspergillus oryzae* TAKA amylase, *Humicola insolens* cellulase, *Humicola insolens* endoglucanase V, *Humicola lanuginosa* lipase, and *Rhizomucor miehei* aspartic proteinase.

[0114] The control sequence may also be a propeptide coding sequence that encodes a propeptide positioned at the N-terminus of a variant. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to an active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding sequence may be obtained

from the genes for *Myceliophthora thermophila* laccase (WO 95/33836), *Rhizomucor miehei* aspartic proteinase.

[0115] Where both signal peptide and propeptide sequences are present, the propeptide sequence is positioned next to the N-terminus of the variant and the signal peptide sequence is positioned next to the N-terminus of the propeptide sequence.

[0116] It may also be desirable to add regulatory sequences that regulate expression of the variant relative to the growth of the host cell. Examples of regulatory systems are those that cause expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems include the *lac, tac,* and *trp* operator systems. In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the *Aspergillus niger* glucoamylase promoter, *Aspergillus oryzae* TAKA alpha-amylase promoter, and *Aspergillus oryzae* glucoamylase promoter may be used. Other examples of regulatory sequences are those that allow for gene amplification. In eukaryotic systems, these regulatory sequences include the dihydrofolate reductase gene that is amplified in the presence of methotrexate, and the metallothionein genes that are amplified with heavy metals. In these cases, the polynucleotide encoding the variant would be operably linked with the regulatory sequence.

Expression Vectors

[0117] The present invention also relates to recombinant expression vectors comprising a polynucleotide encoding a variant of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleotide and control sequences may be joined together to produce a recombinant expression vector that may include one or more convenient restriction sites to allow for insertion or substitution of the polynucleotide encoding the variant at such sites. Alternatively, the polynucleotide may be expressed by inserting the polynucleotide or a nucleic acid construct comprising the polynucleotide into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

[0118] The recombinant expression vector may be any vector (*e.g.,* a plasmid or virus) that can be conveniently subjected to recombinant DNA procedures and can bring about expression of the polynucleotide. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector may be a linear or closed circular plasmid.

[0119] The vector may be an autonomously replicating vector, *i.e.,* a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the

vector may be one that, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids that together contain the total DNA to be introduced into the genome of the host cell, or a transposon, may be used.

[0120] The vector preferably contains one or more selectable markers that permit easy selection of transformed, transfected, transduced, or the like cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

[0121] Selectable markers for use in a filamentous fungal host cell include, but are not limited to, amdS (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hph* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenyltransferase), and *trpC* (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are *Aspergillus nidulans* or *Aspergillus oryzae niaD, niiA, amdS* and *pyrG* genes and a *Streptomyces hygroscopicus bar gene.*

[0122] The vector preferably contains an element(s) that permits integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

[0123] For integration into the host cell genome, the vector may rely on the polynucleotide's sequence encoding the variant or any other element of the vector for integration into the genome by homologous or non-homologous recombination. Alternatively, the vector may contain additional polynucleotides for directing integration by homologous recombination into the genome of the host cell at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should contain a sufficient number of nucleic acids, such as 100 to 10,000 base pairs, 400 to 10,000 base pairs, and 800 to 10,000 base pairs, which have a high degree of sequence identity to the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding polynucleotides. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

[0124] For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. The origin of replication may be any plasmid replicator mediating autonomous replication that functions in a cell. The term "origin of replication" or "plasmid replicator" means a polynucleotide that enables a plasmid or vector to replicate *in vivo.*

[0125] Examples of origins of replication useful in a filamentous fungal cell are AMA1 and ANS1 (Gems et al., 1991, Gene 98: 61-67; Cullen et al., 1987, Nucleic Acids Res. 15: 9163-

9175; WO 00/24883). Isolation of the AMA1 gene and construction of plasmids or vectors comprising the gene can be accomplished according to the methods disclosed in WO 00/24883.

[0126] More than one copy of a polynucleotide of the present invention may be inserted into a host cell to increase production of a variant. An increase in the copy number of the polynucleotide can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the polynucleotide where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the polynucleotide, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

[0127] The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, *e.g.,* Sambrook *et al.,* 1989, *supra).*

Host Cells

[0128] The present invention also relates to recombinant host cells, comprising a polynucleotide encoding a variant of the present invention operably linked to one or more control sequences that direct the production of a variant of the present invention. In a particular embodiment, the recombinant host cell comprises the polynucleotide encoding a trehalase polypeptide of the present invention, wherein the said polynucleotide is heterologous (of different origin/species) to the host cell. A construct or vector comprising a polynucleotide is introduced into a host cell so that the construct or vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication. The choice of a host cell will to a large extent depend upon the gene encoding the variant and its source.

[0129] The host cell may be any cell useful in the recombinant production of a variant, e.g., a prokaryote or a eukaryote.

[0130] The host cell may be a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota as well as the Oomycota and all mitosporic fungi (as defined by Hawksworth et al., In, Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK).

[0131] The fungal host cell may be a yeast cell. "Yeast" as used herein includes ascosporogenous yeast (Endomycetales), basidiosporogenous yeast, and yeast belonging to the Fungi Imperfecti (Blastomycetes). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in Biology and Activities of Yeast (Skinner, Passmore, and Davenport, editors, Soc. App. Bacteriol. Symposium

Series No. 9, 1980).

[0132] The yeast host cell may be a *Candida, Hansenula, Kluyveromyces, Pichia, Saccharomyces, Schizosaccharomyces,* or *Yarrowia* cell such as a *Kluyveromyces lactis, Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norbensis, Saccharomyces oviformis,* or *Yarrowia lipolytica* cell.

[0133] The fungal host cell may be a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth *et al.,* 1995, *supra).* The filamentous fungi are generally characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

[0134] The filamentous fungal host cell may be an *Acremonium, Aspergillus, Aureobasidium, Bjerkandera, Ceriporiopsis, Chrysosporium, Coprinus, Coriolus, Cryptococcus, Filibasidium, Fusarium, Humicola, Magnaporthe, Mucor, Myceliophthora, Neocallimastix, Neurospora, Paecilomyces, Penicillium, Phanerochaete, Phlebia, Piromyces, Pleurotus, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, Trametes,* or *Trichoderma* cell.

[0135] For example, the filamentous fungal host cell may be an *Aspergillus awamori, Aspergillus foetidus, Aspergillus fumigatus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Bjerkandera adusta, Ceriporiopsis aneirina, Ceriporiopsis caregiea, Ceriporiopsis gilvescens, Ceriporiopsis pannocinta, Ceriporiopsis rivulosa, Ceriporiopsis subrufa, Ceriporiopsis subvermispora, Chrysosporium inops, Chrysosporium keratinophilum, Chrysosporium lucknowense, Chrysosporium merdarium, Chrysosporium pannicola, Chrysosporium queenslandicum, Chrysosporium tropicum, Chrysosporium zonatum, Coprinus cinereus, Coriolus hirsutus, Fusarium bactridioides, Fusarium cerealis, Fusarium crookwellense, Fusarium culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium oxysporum, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochroum, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides, Fusarium venenatum, Humicola insolens, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Penicillium purpurogenum, Phanerochaete chrysosporium, Phlebia radiata, Pleurotus eryngii, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei,* or *Trichoderma viride* cell.

[0136] Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se.* Suitable procedures for transformation of *Aspergillus* and *Trichoderma* host cells are described in EP 238023, Yelton et al., 1984, Proc. Natl. Acad. Sci. USA 81: 1470-1474, and Christensen

et al., 1988, BioITechnology6: 1419-1422. Suitable methods for transforming *Fusarium* species are described by Malardier et al., 1989, Gene 78: 147-156, and WO 96/00787. Yeast may be transformed using the procedures described by Becker and Guarente, In Abelson, J.N. and Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito et al., 1983, J. Bacteriol. 153: 163; and Hinnen et al., 1978, Proc. Natl. Acad. Sci. USA 75: 1920.

Methods of Production

[0137] The present invention also relates to methods of producing a variant, comprising: (a) cultivating a recombinant host cell of the present invention under conditions suitable for expression of the variant; and (b) recovering the variant.

[0138] The host cells are cultivated in a nutrient medium suitable for production of the variant using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, or small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the variant to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (*e.g.,* in catalogues of the American Type Culture Collection). If the variant is secreted into the nutrient medium, the variant can be recovered directly from the medium. If the variant is not secreted, it can be recovered from cell lysates.

[0139] The variant may be detected using methods known in the art that are specific for the variants. These detection methods include, but are not limited to, use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme assay may be used to determine the activity of the variant.

[0140] The variant may be recovered using methods known in the art. For example, the variant may be recovered from the nutrient medium by conventional procedures including, but not limited to, collection, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation.

[0141] The variant may be purified by a variety of procedures known in the art including, but not limited to, chromatography (*e.g.,* ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (*e.g.,* preparative isoelectric focusing), differential solubility (*e.g.,* ammonium sulfate precipitation), SDS-PAGE, or extraction (see, *e.g.,* Protein Purification, Janson and Ryden, editors, VCH Publishers, New York, 1989) to obtain substantially pure variants.

[0142] In an alternative aspect, the variant is not recovered, but rather a host cell of the

present invention expressing the variant is used as a source of the variant.

Enzyme Compositions

[0143] The present invention also relates to compositions comprising variant protease of the invention. Preferably, the compositions are enriched in such a protease. The term "enriched" indicates that the pullulanase activity of the composition has been increased, *e.g.,* with an enrichment factor of at least 1.1.

[0144] The compositions may comprise the variant S53 protease as the major enzymatic component, *e.g.,* a mono-component composition. Alternatively, the compositions may comprise multiple enzymatic activities, such as the variant S53 protease and one or more (*e.g.,* several) enzymes selected from the group consisting of hydrolase, isomerase, ligase, lyase, oxidoreductase, or transferase, *e.g.,* an alpha-galactosidase, alpha-glucosidase, aminopeptidase, alpha-amylase, beta-amylase, beta-galactosidase, beta-glucosidase, beta-xylosidase, carbohydrase, carboxypeptidase, catalase, cellobiohydrolase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, endoglucanase, esterase, glucoamylase, invertase, laccase, lipase, mannosidase, mutanase, oxidase, pectinolytic enzyme, peroxidase, phytase, polyphenoloxidase, protease, ribonuclease, transglutaminase, or xylanase. In one embodiment the composition comprises a variant S53 protease of the invention and a carbohydrate-source generating enzyme and optionally an alpha-amylase. In one particular embodiment the composition comprises a variant S53 protease and a glucoamylase. Preferably the enzyme activities comprised in the composition are selected from the variant S53 protease of the invention and one or more enzymes selected from the group consisting of glucoamylase, fungal alpha-amylase.

[0145] In an embodiment the glucoamylase comprised in the composition is of fungal origin, preferably from a stain of *Aspergillus,* preferably *A. niger, A. awamori,* or *A. oryzae;* or a strain of *Trichoderma,* preferably *T. reesei;* or a strain of *Talaromyces,* preferably *T. emersonii* or a strain of *Trametes,* preferably *T. cingulata,* or a strain of *Pycnoporus,* preferable *P. sanguineus,* or a strain of *Gloeophyllum,* such as *G. serpiarium* or *G. trabeum,* or a strain of the *Nigrofomes.*

[0146] In an embodiment the glucoamylase is derived from *Trametes,* such as a strain of *Trametes cingulata,* such as the one shown in SEQ ID NO: 4 herein.

[0147] In an embodiment the glucoamylase is selected from the group consisting of:

1. (i) a glucoamylase comprising the polypeptide of SEQ ID NO: 4 herein;
2. (ii) a glucoamylase comprising an amino acid sequence having at least 60%, at least 70%, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the polypeptide of SEQ ID NO: 4 herein.

[0148] In an embodiment the glucoamylase is derived from *Talaromyces,* such as a strain of *Talaromyces emersonii,* such as the one shown in SEQ ID NO: 5 herein.

[0149] In an embodiment the glucoamylase is selected from the group consisting of:

1. (i) a glucoamylase comprising the polypeptide of SEQ ID NO: 5 herein;
2. (ii) a glucoamylase comprising an amino acid sequence having at least 60%, at least 70%, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the polypeptide of SEQ ID NO: 5 herein.

[0150] In an embodiment the glucoamylase is derived from a strain of the genus *Pycnoporus,* in particular a strain of *Pycnoporus sanguineus* described in WO 2011/066576 (SEQ ID NOs 2, 4 or 6), such as the one shown as SEQ ID NO: 4 in WO 2011/066576.

[0151] In an embodiment the glucoamylase is selected from the group consisting of:

1. (i) a glucoamylase comprising the polypeptide of SEQ ID NO: 6 herein;
2. (ii) a glucoamylase comprising an amino acid sequence having at least 60%, at least 70%, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at SEQ ID NO: 6 herein.

[0152] In an embodiment the glucoamylase is derived from a strain of the genus *Gloeophyllum,* such as a strain of *Gloeophyllum sepiarium* or *Gloeophyllum trabeum,* in particular a strain of *Gloeophyllum* as described in WO 2011/068803 (SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 or 16). In a preferred embodiment the glucoamylase is the *Gloeophyllum sepiarium* shown in SEQ ID NO: 2 in WO 2011/068803 or SEQ ID NO: 7 herein.

[0153] In an embodiment the glucoamylase is derived from *Gloeophyllum serpiarium,* such as the one shown in SEQ ID NO: 7 herein. In an embodiment the glucoamylase is selected from the group consisting of:

1. (i) a glucoamylase comprising the polypeptide of SEQ ID NO: 7 herein;
2. (ii) a glucoamylase comprising an amino acid sequence having at least 60%, at least 70%, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the polypeptide of SEQ ID NO: 7 herein.

[0154] In another embodiment the glucoamylase is derived from *Gloeophyllum trabeum* such as the one shown in SEQ ID NO: 8 herein. In an embodiment the glucoamylase is selected from the group consisting of:

1. (i) a glucoamylase comprising the polypeptide of SEQ ID NO: 8 herein;
2. (ii) a glucoamylase comprising an amino acid sequence having at least 60%, at least 70%, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the polypeptide of SEQ ID NO: 8 herein.

[0155] In an embodiment the glucoamylase is derived from a strain of the genus *Nigrofomes,* in particular a strain of *Nigrofomes sp.* disclosed in WO 2012/064351.

[0156] Glucoamylases may in an embodiment be added to the saccharification and/or fermentation in an amount of 0.0001-20 AGU/g DS, preferably 0.001-10 AGU/g DS, especially between 0.01-5 AGU/g DS, such as 0.1-2 AGU/g DS.

[0157] Commercially available compositions comprising glucoamylase include AMG 200L; AMG 300 L; SAN™ SUPER, SAN™ EXTRA L, SPIRIZYME™ PLUS, SPIRIZYME™ FUEL, SPIRIZYME™ B4U, SPIRIZYME™ ULTRA, SPIRIZYME™ EXCEL and AMG™ E (from Novozymes A/S); OPTIDEX™ 300, GC480, GC417 (from DuPont.); AMIGASE™ and AMIGASE™ PLUS (from DSM); G-ZYME™ G900, G-ZYME™ and G990 ZR (from DuPont).

[0158] In addition to a glucoamylase the composition may further comprise an alpha-amylase. Particularly the alpha-amylase is an acid fungal alpha-amylase. A fungal acid stable alpha-amylase is an alpha-amylase that has activity in the pH range of 3.0 to 7.0 and preferably in the pH range from 3.5 to 6.5, including activity at a pH of about 4.0, 4.5, 5.0, 5.5, and 6.0.

[0159] Preferably the acid fungal alpha-amylase is derived from the genus *Aspergillus,* especially a strain of *A. terreus, A. niger, A. oryzae,* A. *awamori,* or *Aspergillus kawachii,* or from the genus *Rhizomucor,* preferably a strain the *Rhizomucor pusillus,* or the genus *Meripilus,* preferably a strain of *Meripilus giganteus.*

[0160] In a preferred embodiment the alpha-amylase is derived from a strain of the genus *Rhizomucor,* preferably a strain the *Rhizomucorpusillus,* such as one shown in SEQ ID NO: 3 in WO 2013/006756, such as a *Rhizomucor pusillus* alpha-amylase hybrid having an *Aspergillus niger* linker and starch-binding domain, such as the one shown in SEQ ID NO: 9 herein, or a variant thereof.

[0161] In an embodiment the alpha-amylase is selected from the group consisting of:

1. (i) an alpha-amylase comprising the polypeptide of SEQ ID NO: 9 herein;

2. (ii) an alpha-amylase comprising an amino acid sequence having at least 60%, at least 70%, *e.g.,* at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the polypeptide of SEQ ID NO: 9 herein.

[0162] In a preferred embodiment the alpha-amylase is a variant of the alpha-amylase shown in SEQ ID NO: 9 having at least one of the following substitutions or combinations of substitutions: D165M; Y141W; Y141R; K136F; K192R; P224A; P224R; S123H + Y141W; G20S + Y141W; A76G + Y141W; G128D + Y141W; G128D + D143N; P219C + Y141W; N142D + D143N; Y141W + K192R; Y141W + D143N; Y141W + N383R; Y141W + P219C + A265C; Y141W + N142D + D143N; Y141W + K192R V410A; G128D + Y141W + D143N; Y141W + D143N + P219C; Y141W + D143N + K192R; G128D + D143N + K192R; Y141W + D143N + K192R + P219C; G128D + Y141W + D143N + K192R; or G128D + Y141W + D143N + K192R + P219C (using SEQ ID NO: 9 for numbering).

[0163] In an embodiment the alpha-amylase is derived from a *Rhizomucor pusillus* with an *Aspergillus niger* glucoamylase linker and starch-binding domain (SBD), preferably disclosed as SEQ ID NO: 9 herein, preferably having one or more of the following substitutions: G128D, D143N, preferably G128D+D143N (using SEQ ID NO: 9 for numbering), and wherein the alpha-amylase variant has at least 75% identity preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, even more preferably at least 93%, most preferably at least 94%, and even most preferably at least 95%, such as even at least 96%, at least 97%, at least 98%, at least 99%, but less than 100% identity to the polypeptide of SEQ ID NO: 9 herein.

[0164] In a preferred embodiment the ratio between glucoamylase and alpha-amylase present and/or added during saccharification and/or fermentation may preferably be in the range from 500:1 to 1:1, such as from 250:1 to 1:1, such as from 100:1 to 1: 1, such as from 100: 2 to 100:50, such as from 100:3 to 100:70.

[0165] The compositions may be prepared in accordance with methods known in the art and may be in the form of a liquid or a dry composition. For instance, the composition may be in the form of granulate or microgranulate. The variant may be stabilized in accordance with methods known in the art.

[0166] The compositions may be prepared in accordance with methods known in the art and may be in the form of a liquid or a dry composition. The compositions may be stabilized in accordance with methods known in the art.

[0167] The enzyme composition of the present invention may be in any form suitable for use, such as, for example, a crude fermentation broth with or without cells removed, a cell lysate with or without cellular debris, a semi-purified or purified enzyme composition, or a host cell, as

a source of the enzymes.

[0168] The enzyme composition may be a dry powder or granulate, a non-dusting granulate, a liquid, a stabilized liquid, or a stabilized protected enzyme. Liquid enzyme compositions may, for instance, be stabilized by adding stabilizers such as a sugar, a sugar alcohol or another polyol, and/or lactic acid or another organic acid according to established processes.

Use of the variant proteases of the invention

Starch Processing

[0169] Native starch consists of microscopic granules, which are insoluble in water at room temperature. When aqueous starch slurry is heated, the granules swell and eventually burst, dispersing the starch molecules into the solution. At temperatures up to about 50°C to 75°C the swelling may be reversible. However, with higher temperatures an irreversible swelling called "gelatinization" begins. During this "gelatinization" process there is a dramatic increase in viscosity. Granular starch to be processed may be a highly refined starch quality, preferably at least 90%, at least 95%, at least 97% or at least 99.5% pure or it may be a more crude starch-containing materials comprising (*e.g.,* milled) whole grains including non-starch fractions such as germ residues and fibers. The raw material, such as whole grains, may be reduced in particle size, e.g., by milling, in order to open up the structure and allowing for further processing. In dry milling whole kernels are milled and used. Wet milling gives a good separation of germ and meal (starch granules and protein) and is often applied at locations where the starch hydrolysate is used in the production of, *e.g.,* syrups. Both dry and wet milling is well known in the art of starch processing and may be used in a process of the invention. Methods for reducing the particle size of the starch containing material are well known to those skilled in the art.

[0170] As the solids level is 30-40% in a typical industrial process, the starch has to be thinned or "liquefied" so that it can be suitably processed. This reduction in viscosity is primarily attained by enzymatic degradation in current commercial practice.

[0171] Liquefaction is carried out in the presence of an alpha-amylase, preferably a bacterial alpha-amylase and/or acid fungal alpha-amylase. In an embodiment, a phytase is also present during liquefaction. In an embodiment, viscosity reducing enzymes such as a xylanase and/or beta-glucanase is also present during liquefaction.

[0172] During liquefaction, the long-chained starch is degraded into branched and linear shorter units (maltodextrins) by an alpha-amylase. Liquefaction may be carried out as a three-step hot slurry process. The slurry is heated to between 60-95°C (*e.g.,* 70-90°C, such as 77-86°C, 80-85°C, 83-85°C) and an alpha-amylase is added to initiate liquefaction (thinning).

[0173] The slurry may in an embodiment be jet-cooked at between 95-140°C, *e.g.,* 105-125°C, for about 1-15 minutes, *e.g.,* about 3-10 minutes, especially around 5 minutes. The slurry is then cooled to 60-95°C and more alpha-amylase is added to obtain final hydrolysis (secondary liquefaction). The jet-cooking process is carried out at pH 4.5-6.5, typically at a pH between 5 and 6. The alpha-amylase may be added as a single dose, *e.g.,* before jet cooking.

[0174] The liquefaction process is carried out at between 70-95°C, such as 80-90°C, such as around 85°C, for about 10 minutes to 5 hours, typically for 1-2 hours. The pH is between 4 and 7, such as between 4.5 and 5.5. In order to ensure optimal enzyme stability under these conditions, calcium may optionally be added (to provide 1-60 ppm free calcium ions, such as about 40 ppm free calcium ions). After such treatment, the liquefied starch will typically have a "dextrose equivalent" (DE) of 10-15.

[0175] Generally liquefaction and liquefaction conditions are well known in the art.

[0176] Alpha-amylases for use in liquefaction are preferably bacterial acid stable alpha-amylases. Particularly the alpha-amylase is from an *Exiguobacterium* sp. or a *Bacillus* sp. such as e.g., *Bacillus stearothermophilus* or *Bacillus licheniformis.*

[0177] In an embodiment the alpha-amylase is from the genus *Bacillus,* such as a strain of *Bacillus stearothermophilus,* in particular a variant of a *Bacillus stearothermophilus* alpha-amylase, such as the one shown in SEQ ID NO: 3 in WO 99/019467 or SEQ ID NO: 10 herein.

[0178] In an embodiment the *Bacillus stearothermophilus* alpha-amylase has a double deletion of two amino acids in the region from position 179 to 182, more particularly a double deletion at positions I181 + G182, R179 + G180, G180 + I181, R179 + I181, or G180 + G182, preferably I181 + G182, and optionally a N193F substitution, (using SEQ ID NO: 10 for numbering).

[0179] In an embodiment the *Bacillus stearothermophilus* alpha-amylase has a substitution at position S242, preferably S242Q substitution.

[0180] In an embodiment the *Bacillus stearothermophilus* alpha-amylase has a substitution at position E188, preferably E188P substitution.

[0181] In an embodiment the alpha-amylase is selected from the group of *Bacillus stearothermophilus* alpha-amylase variants with the following mutations:

- I181*+G182*+N193F+E129V+K177L+R179E;
- I181*+G182*+N193F+V59A+Q89R+E129V+K177L+R179E+H208Y+K220P+N224L +Q254S;
- I181*+G182*+N193F +V59A Q89R+ E129V+ K177L+ R179E+ Q254S+ M284V; and
- I181*+G182*+N193F+E129V+K177L+R179E+K220P+N224L+S242Q+Q254S (using SEQ ID NO: 10 for numbering).

[0182] In an embodiment the alpha-amylase variant has at least 75% identity preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, even more preferably at least 93%, most preferably at least 94%, and even most preferably at least 95%, such as even at least 96%, at least 97%, at least 98%, at least 99%, but less than 100% identity to the polypeptide of SEQ ID NO: 10.

[0183] It should be understood that when referring to *Bacillus stearothermophilus* alpha-amylase and variants thereof they are normally produced in truncated form. In particular, the truncation may be so that the *Bacillus stearothermophilus* alpha-amylase shown in SEQ ID NO: 3 in WO 99/19467 or SEQ ID NO: 10 herein, or variants thereof, are truncated in the C-terminal preferably to have around 490 amino acids, such as from 482-493 amino acids. Preferably the *Bacillus stearothermophilus* variant alpha-amylase is truncated, preferably after position 484 of SEQ ID NO: 10, particularly after position 485, particularly after position 486, particularly after position 487, particularly after position 488, particularly after position 489, particularly after position 490, particularly after position 491, particularly after position 492, more particularly after position 493.

[0184] Saccharification may be carried out using conditions well-known in the art with a carbohydrate-source generating enzyme, in particular a glucoamylase, or a beta-amylase and optionally a debranching enzyme, such as an isoamylase or a pullulanase. For instance, a full saccharification step may last from about 24 to about 72 hours. However, it is common to do a pre-saccharification of typically 40-90 minutes at a temperature between 30-65°C, typically about 60°C, followed by complete saccharification during fermentation in a simultaneous saccharification and fermentation (SSF) process. Saccharification is typically carried out at a temperature in the range of 20-75°C, *e.g.,* 25-65°C and 40-70°C, typically around 60°C, and at a pH between about 4 and 5, normally at about pH 4.5.

[0185] The saccharification and fermentation steps may be carried out either sequentially or simultaneously. In an embodiment, saccharification and fermentation are performed simultaneously (referred to as "SSF"). However, it is common to perform a pre-saccharification step for about 30 minutes to 2 hours (*e.g.,* 30 to 90 minutes) at a temperature of 30 to 65°C, typically around 60°C which is followed by a complete saccharification during fermentation referred to as simultaneous saccharification and fermentation (SSF). The pH is usually between 4.2-4.8, *e.g.,* pH 4.5. In a simultaneous saccharification and fermentation (SSF) process, there is no holding stage for saccharification, rather, the yeast and enzymes are added together.

[0186] In a typical saccharification process, maltodextrins produced during liquefaction are converted into dextrose by adding a glucoamylase and a debranching enzyme, such as an isoamylase (U.S. Patent No. 4,335,208) or a pullulanase. The temperature is lowered to 60°C, prior to the addition of the glucoamylase and debranching enzyme. The saccharification process proceeds for 24-72 hours. Prior to addition of the saccharifying enzymes, the pH is

reduced to below 4.5, while maintaining a high temperature (above 95°C), to inactivate the liquefying alpha-amylase. This process reduces the formation of short oligosaccharide called "panose precursors," which cannot be hydrolyzed properly by the debranching enzyme. Normally, about 0.2-0.5% of the saccharification product is the branched trisaccharide panose (Glc pα1-6Glc pα1-4Glc), which cannot be degraded by a pullulanase. If active amylase from the liquefaction remains present during saccharification (*i.e.,* no denaturing), the amount of panose can be as high as 1-2%, which is highly undesirable since it lowers the saccharification yield significantly.

[0187] Other fermentation products may be fermented at conditions and temperatures well known to persons skilled in the art, suitable for the fermenting organism in question.

[0188] The fermentation product may be recovered by methods well known in the art, e.g., by distillation. Examples of carbohydrate-source generating enzymes are disclosed in the "Enzymes" section below.

[0189] In a particular embodiment, the process of the invention further comprises, prior to the conversion of a starch-containing material to sugars/dextrins the steps of:

(x) reducing the particle size of the starch-containing material; and

(y) forming a slurry comprising the starch-containing material and water.

[0190] In an embodiment, the starch-containing material is milled to reduce the particle size. In an embodiment the particle size is reduced to between 0.05-3.0 mm, preferably 0.1-0.5 mm, or so that at least 30%, preferably at least 50%, more preferably at least 70%, even more preferably at least 90% of the starch-containing material fits through a sieve with a 0.05-3.0 mm screen, preferably 0.1-0.5 mm screen.

[0191] The aqueous slurry may contain from 10-55 wt. % dry solids (DS), preferably 25-45 wt. % dry solids (DS), more preferably 30-40 wt. % dry solids (DS) of starch-containing material.

[0192] Conventional starch-conversion processes, such as liquefaction and saccharification processes are described, e.g., in U.S. Patent No. 3,912,590, EP 252730 and EP 063909.

[0193] In an embodiment, the conversion process degrading starch to lower molecular weight carbohydrate components such as sugars or fat replacers includes a debranching step.

[0194] In the case of converting starch into a sugar, the starch is depolymerized. Such a depolymerization process consists of, *e.g.,* a pre-treatment step and two or three consecutive process steps, *i.e.,* a liquefaction process, a saccharification process, and depending on the desired end-product, an optional isomerization process.

[0195] When the desired final sugar product is, *e.g.,* high fructose syrup the dextrose syrup may be converted into fructose. After the saccharification process, the pH is increased to a value in the range of 6-8, *e.g.,* pH 7.5, and the calcium is removed by ion exchange. The dextrose syrup is then converted into high fructose syrup using, *e.g.,* an immobilized glucose isomerase.

Production of Fermentation Products

[0196] Fermentable sugars (*e.g.*, dextrins, monosaccharides, particularly glucose) are produced from enzymatic saccharification. These fermentable sugars may be further purified and/or converted to useful sugar products. In addition, the sugars may be used as a fermentation feedstock in a microbial fermentation process for producing end-products, such as alcohol (*e.g.,* ethanol, and butanol), organic acids (*e.g.,* succinic acid, 3-HP and lactic acid), sugar alcohols (*e.g.,* glycerol), ascorbic acid intermediates (*e.g.,* gluconate, 2-keto-D-gluconate, 2,5-diketo-D-gluconate, and 2-keto-L-gulonic acid), amino acids (*e.g.,* lysine), proteins (*e.g.,* antibodies and fragment thereof).

[0197] In an embodiment, the fermentable sugars obtained during the liquefaction process steps are used to produce alcohol and particularly ethanol. In ethanol production, an SSF process is commonly used wherein the saccharifying enzymes and fermenting organisms (*e.g.,* yeast) are added together and then carried out at a temperature of 30-40°C.

[0198] The organism used in fermentation will depend on the desired end-product. Typically, if ethanol is the desired end product yeast will be used as the fermenting organism. In some preferred embodiments, the ethanol-producing microorganism is a yeast and specifically *Saccharomyces* such as strains of *S. cerevisiae* (U.S. Patent No. 4,316,956). A variety of S. *cerevisiae* are commercially available and these include but are not limited to FALI (Fleischmann's Yeast), SUPERSTART (Alltech), FERMIOL (DSM Specialties), RED STAR (Lesaffre) and Angel alcohol yeast (Angel Yeast Company, China). The amount of starter yeast employed in the methods is an amount effective to produce a commercially significant amount of ethanol in a suitable amount of time, (*e.g.,* to produce at least 10% ethanol from a substrate having between 25-40% DS in less than 72 hours). Yeast cells are generally supplied in amounts of about 10^4 to about 10^{12}, and preferably from about 10^7 to about 10^{10} viable yeast count per mL of fermentation broth. After yeast is added to the mash, it is typically subjected to fermentation for about 24-96 hours, e.g., 35-60 hours. The temperature is between about 26-34°C, typically at about 32°C, and the pH is from pH 3-6, e.g., around pH 4-5.

[0199] The fermentation may include, in addition to a fermenting microorganisms (*e.g.,* yeast), nutrients, and additional enzymes, including phytases. The use of yeast in fermentation is well known in the art.

[0200] In further embodiments, use of appropriate fermenting microorganisms, as is known in

the art, can result in fermentation end product including, e.g., glycerol, 1,3-propanediol, gluconate, 2-keto-D-gluconate, 2,5-diketo-D-gluconate, 2-keto-L-gulonic acid, succinic acid, lactic acid, amino acids, and derivatives thereof. More specifically when lactic acid is the desired end product, a *Lactobacillus* sp. (*L. casei*) may be used; when glycerol or 1,3-propanediol are the desired end-products *E. coli* may be used; and when 2-keto-D-gluconate, 2,5-diketo-D-gluconate, and 2-keto-L-gulonic acid are the desired end products, *Pantoea citrea* may be used as the fermenting microorganism. The above enumerated list are only examples and one skilled in the art will be aware of a number of fermenting microorganisms that may be used to obtain a desired end product.

Processes for producing fermentation products from un-gelatinized starch-containing material

[0201] The invention relates to processes for producing fermentation products from starch-containing material without gelatinization (*i.e.,* without cooking) of the starch-containing material (often referred to as a "raw starch hydrolysis" process). The fermentation product, such as ethanol, can be produced without liquefying the aqueous slurry containing the starch-containing material and water. In one embodiment a process of the invention includes saccharifying (*e.g.,* milled) starch-containing material, *e.g.,* granular starch, below the initial gelatinization temperature, preferably in the presence of alpha-amylase and/or carbohydrate-source generating enzyme(s) to produce sugars that can be fermented into the fermentation product by a suitable fermenting organism. In this embodiment the desired fermentation product, *e.g.,* ethanol, is produced from un-gelatinized (*i.e.*, uncooked), preferably milled, cereal grains, such as corn.

[0202] Accordingly, in one aspect the invention relates to processes for producing a fermentation product from starch-containing material comprising simultaneously saccharifying and fermenting starch-containing material using a carbohydrate-source generating enzymes and a fermenting organism at a temperature below the initial gelatinization temperature of said starch-containing material in the presence of a variant protease of the invention. Saccharification and fermentation may also be separate. Thus in another aspect the invention relates to processes of producing fermentation products, comprising the following steps:

1. (i) saccharifying a starch-containing material at a temperature below the initial gelatinization temperature using a carbohydrate-source generating enzyme, e.g., a glucoamylase; and
2. (ii) fermenting using a fermentation organism;

wherein step (i) is carried out using at least a glucoamylase, and a variant protease of the invention.

[0203] In one embodiment the fermenting organism expresses the variant protease of the invention.

[0204] In one embodiment, an alpha amylase is also added in step (i). Steps (i) and (ii) may be performed simultaneously.

[0205] The fermentation product, *e.g.,* ethanol, may optionally be recovered after fermentation, *e.g.,* by distillation. Typically amylase(s), such as glucoamylase(s) and/or other carbohydrate-source generating enzymes, and/or alpha-amylase(s), is(are) present during fermentation. Examples of glucoamylases and other carbohydrate-source generating enzymes include raw starch hydrolyzing glucoamylases. Examples of alpha-amylase(s) include acid alpha-amylases such as acid fungal alpha-amylases. Examples of fermenting organisms include yeast, *e.g.,* a strain of *Saccharomyces cerevisiae.* The term "initial gelatinization temperature" means the lowest temperature at which starch gelatinization commences. In general, starch heated in water begins to gelatinize between about 50°C and 75°C; the exact temperature of gelatinization depends on the specific starch and can readily be determined by the skilled artisan. Thus, the initial gelatinization temperature may vary according to the plant species, to the particular variety of the plant species as well as with the growth conditions. In the context of this invention the initial gelatinization temperature of a given starch-containing material may be determined as the temperature at which birefringence is lost in 5% of the starch granules using the method described by Gorinstein and Lii, 1992, Starch/Stärke 44(12): 461-466. Before initiating the process a slurry of starch-containing material, such as granular starch, having 10-55 w/w % dry solids (DS), preferably 25-45 w/w % dry solids, more preferably 30-40 w/w % dry solids of starch-containing material may be prepared. The slurry may include water and/or process waters, such as stillage (backset), scrubber water, evaporator condensate or distillate, side-stripper water from distillation, or process water from other fermentation product plants. Because the process of the invention is carried out below the initial gelatinization temperature, and thus no significant viscosity increase takes place, high levels of stillage may be used if desired. In an embodiment the aqueous slurry contains from about 1 to about 70 vol. %, preferably 15-60 vol. %, especially from about 30 to 50 vol. % water and/or process waters, such as stillage (backset), scrubber water, evaporator condensate or distillate, side-stripper water from distillation, or process water from other fermentation product plants, or combinations thereof, or the like. The starch-containing material may be prepared by reducing the particle size, preferably by dry or wet milling, to 0.05 to 3.0 mm, preferably 0.1-0.5 mm. After being subjected to a process of the invention at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or preferably at least 99% of the dry solids in the starch-containing material are converted into a soluble starch hydrolyzate. A process in this aspect of the invention is conducted at a temperature below the initial gelatinization temperature, which means that the temperature typically lies in the range between 30-75°C, preferably between 45-60°C. In a preferred embodiment the process carried at a temperature from 25°C to 40°C, such as from 28°C to 35°C, such as from 30°C to 34°C, preferably around 32°C. In an embodiment the process is carried out so that the sugar level, such as glucose level, is kept at a low level, such as below 6 w/w %, such as below about 3 w/w %, such as below about 2 w/w %, such as below about 1 w/w %., such as below about 0.5 w/w %, or below 0.25 w/w %, such as below about 0.1 w/w %. Such low levels of sugar can be accomplished by simply employing adjusted quantities of enzyme and fermenting organism.

A skilled person in the art can easily determine which doses/quantities of enzyme and fermenting organism to use. The employed quantities of enzyme and fermenting organism may also be selected to maintain low concentrations of maltose in the fermentation broth. For instance, the maltose level may be kept below about 0.5 w/w %, such as below about 0.2 w/w %. The process of the invention may be carried out at a pH from about 3 and 7, preferably from pH 3.5 to 6, or more preferably from pH 4 to 5. In an embodiment fermentation is ongoing for 6 to 120 hours, in particular 24 to 96 hours.

Processes for producing fermentation products from gelatinized starch-containing material

[0206] In this aspect, the invention relates to processes for producing fermentation products, especially ethanol, from starch-containing material, which process includes a liquefaction step and sequentially or simultaneously performed saccharification and fermentation steps. Consequently, the invention relates to a process for producing a fermentation product from starch-containing material comprising the steps of:

1. (a) liquefying starch-containing material in the presence of an alpha-amylase;
2. (b) saccharifying the liquefied material obtained in step (a) using a carbohydrate-source generating enzyme;
3. (c) fermenting using a fermenting organism;

wherein a variant protease of the invention is present during step b) and/or c).

[0207] In one embodiment, the fermenting organism expresses the variant protease of the invention.

[0208] The fermentation product, such as especially ethanol, may optionally be recovered after fermentation, e.g., by distillation. The fermenting organism is preferably yeast, preferably a strain of *Saccharomyces cerevisiae.* In a particular embodiment, the process of the invention further comprises, prior to step (a), the steps of:

x) reducing the particle size of the starch-containing material, preferably by milling (e.g., using a hammer mill);

y) forming a slurry comprising the starch-containing material and water.

[0209] In an embodiment, the particle size is smaller than a # 7 screen, *e.g.,* a # 6 screen. A # 7 screen is usually used in conventional prior art processes. The aqueous slurry may contain from 10-55, *e.g.,* 25-45 and 30-40, w/w % dry solids (DS) of starch-containing material. The slurry is heated to above the gelatinization temperature and an alpha-amylase variant may be added to initiate liquefaction (thinning). The slurry may in an embodiment be jet-cooked to

further gelatinize the slurry before being subjected to alpha-amylase in step (a). Liquefaction may in an embodiment be carried out as a three-step hot slurry process. The slurry is heated to between 60-95°C, preferably between 70-90°C, such as preferably between 80-85°Cat pH 4-6, preferably 4.5-5.5, and alpha-amylase variant, optionally together with a pullulanase and/or protease, preferably metalloprotease, are added to initiate liquefaction (thinning). In an embodiment the slurry may then be jet-cooked at a temperature between 95-140°C, preferably 100-135°C, such as 105-125°C, for about 1-15 minutes, preferably for about 3-10 minutes, especially around about 5 minutes. The slurry is cooled to 60-95°C and more alpha-amylase variant and optionally pullulanase variant and/or protease, preferably metalloprotease, is(are) added to finalize hydrolysis (secondary liquefaction). The liquefaction process is usually carried out at pH 4.0-6, in particular at a pH from 4.5 to 5.5. Saccharification step (b) may be carried out using conditions well known in the art. For instance, a full saccharification process may last up to from about 24 to about 72 hours, however, it is common only to do a pre-saccharification of typically 40-90 minutes at a temperature between 30-65°C, typically about 60°C, followed by complete saccharification during fermentation in a simultaneous saccharification and fermentation process (SSF process). Saccharification is typically carried out at temperatures from 20-75°C, preferably from 40-70°C, typically around 60°C, and at a pH between 4 and 5, normally at about pH 4.5. The most widely used process to produce a fermentation product, especially ethanol, is a simultaneous saccharification and fermentation (SSF) process, in which there is no holding stage for the saccharification, meaning that a fermenting organism, such as yeast, and enzyme(s), may be added together. SSF may typically be carried out at a temperature from 25°C to 40°C, such as from 28°C to 35°C, such as from 30°C to 34°C, preferably around about 32°C. In an embodiment fermentation is ongoing for 6 to 120 hours, in particular 24 to 96 hours. Glucoamylase Present And/Or Added In Saccharification And/Or Fermentation

[0210] The carbohydrate-source generating enzyme present during saccharification may in one embodiment be a glucoamylase. A glucoamylase is present and/or added in saccharification and/or fermentation, preferably simultaneous saccharification and fermentation (SSF), in a process of the invention (i.e., saccharification and fermentation of ungelatinized or gelatinized starch material).

[0211] In an embodiment the glucoamylase present and/or added in saccharification and/or fermentation is of fungal origin, preferably from a stain of *Aspergillus,* preferably *A. niger, A. awamori,* or *A. oryzae;* or a strain of *Trichoderma,* preferably *T. reesei;* or a strain of *Talaromyces,* preferably *T. emersonii* or a strain of *Trametes,* preferably *T. cingulata,* or a strain of *Pycnoporus,* preferably *P. sanguineus,* or a strain of *Gloeophyllum,* such as *G. serpiarium* or *G. trabeum,* or a strain of the *Nigrofomes.*

[0212] In an embodiment the glucoamylase is derived from *Trametes,* such as a strain of *Trametes cingulata,* such as the one shown in SEQ ID NO: 4 herein.

[0213] In an embodiment the glucoamylase is selected from the group consisting of:

1. (i) a glucoamylase comprising the polypeptide of SEQ ID NO: 4 herein;
2. (ii) a glucoamylase comprising an amino acid sequence having at least 60%, at least 70%, *e.g.,* at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the polypeptide of SEQ ID NO: 4 herein.

[0214] In an embodiment the glucoamylase is derived from *Talaromyces,* such as a strain of *Talaromyces emersonii,* such as the one shown in SEQ ID NO: 5 herein.

[0215] In an embodiment the glucoamylase is selected from the group consisting of:

1. (i) a glucoamylase comprising the polypeptide of SEQ ID NO: 5 herein;
2. (ii) a glucoamylase comprising an amino acid sequence having at least 60%, at least 70%, *e.g.,* at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the polypeptide of SEQ ID NO: 5 herein.

[0216] In an embodiment the glucoamylase is derived from a strain of the genus *Pycnoporus,* in particular a strain of *Pycnoporus sanguineus* described in WO 2011/066576 (SEQ ID NOs 2, 4 or 6), such as the one shown as SEQ ID NO: 4 in WO 2011/066576.

[0217] In an embodiment the glucoamylase is selected from the group consisting of:

1. (i) a glucoamylase comprising the polypeptide of SEQ ID NO: 6 herein;
2. (ii) a glucoamylase comprising an amino acid sequence having at least 60%, at least 70%, *e.g.,* at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at 6 herein.

[0218] In an embodiment the glucoamylase is derived from a strain of the genus *Gloeophyllum,* such as a strain of *Gloeophyllum sepiarium* or *Gloeophyllum trabeum,* in particular a strain of *Gloeophyllum* as described in WO 2011/068803 (SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 or 16). In a preferred embodiment the glucoamylase is the *Gloeophyllum sepiarium* shown in SEQ ID NO: 2 in WO 2011/068803 or SEQ ID NO: 7 herein.

[0219] In an embodiment the glucoamylase is derived from *Gloeophyllum serpiarium,* such as the one shown in SEQ ID NO: 7 herein. In an embodiment the glucoamylase is selected from the group consisting of:

1. (i) a glucoamylase comprising the polypeptide of SEQ ID NO: 7 herein;
2. (ii) a glucoamylase comprising an amino acid sequence having at least 60%, at least

70%, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the polypeptide of SEQ ID NO: 7 herein.

[0220] In another embodiment the glucoamylase is derived from *Gloeophyllum trabeum* such as the one shown in SEQ ID NO: 8 herein. In an embodiment the glucoamylase is selected from the group consisting of:

1. (i) a glucoamylase comprising the polypeptide of SEQ ID NO: 8 herein;
2. (ii) a glucoamylase comprising an amino acid sequence having at least 60%, at least 70%, *e.g.,* at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the polypeptide of SEQ ID NO: 8 herein.

[0221] In an embodiment the glucoamylase is derived from a strain of the genus *Nigrofomes,* in particular a strain of *Nigrofomes sp.* disclosed in WO 2012/064351.

[0222] Glucoamylases may in an embodiment be added to the saccharification and/or fermentation in an amount of 0.0001-20 AGU/g DS, preferably 0.001-10 AGU/g DS, especially between 0.01-5 AGU/g DS, such as 0.1-2 AGU/g DS.

[0223] Commercially available compositions comprising glucoamylase include AMG 200L; AMG 300 L; SAN™ SUPER, SAN™ EXTRA L, SPIRIZYME™ PLUS, SPIRIZYME™ FUEL, SPIRIZYME™ B4U, SPIRIZYME™ ULTRA, SPIRIZYME™ EXCEL and AMG™ E (from Novozymes A/S); OPTIDEX™ 300, GC480, GC417 (from DuPont.); AMIGASE™ and AMIGASE™ PLUS (from DSM); G-ZYME™ G900, G-ZYME™ and G990 ZR (from DuPont).

[0224] According to a preferred embodiment of the invention the glucoamylase is present and/or added in saccharification and/or fermentation in combination with an alpha-amylase. Examples of suitable alpha-amylase are described below.

Alpha-Amylase Present and/or Added In Saccharification And/Or Fermentation

[0225] In an embodiment an alpha-amylase is present and/or added in saccharification and/or fermentation in the processes of the invention. In a preferred embodiment the alpha-amylase is of fungal or bacterial origin. In a preferred embodiment the alpha-amylase is a fungal acid stable alpha-amylase. A fungal acid stable alpha-amylase is an alpha-amylase that has activity in the pH range of 3.0 to 7.0 and preferably in the pH range from 3.5 to 6.5, including activity at a pH of about 4.0, 4.5, 5.0, 5.5, and 6.0.

[0226] In a preferred embodiment the alpha-amylase present and/or added in saccharification and/or fermentation is derived from a strain of the genus *Rhizomucor,* preferably a strain the *Rhizomucor pusillus,* such as one shown in SEQ ID NO: 3 in WO 2013/006756, such as a *Rhizomucor pusillus* alpha-amylase hybrid having an *Aspergillus niger* linker and starch-binding domain, such as the one shown in SEQ ID NO: 9 herein, or a variant thereof.

[0227] In an embodiment the alpha-amylase present and/or added in saccharification and/or fermentation is selected from the group consisting of:

1. (i) an alpha-amylase comprising the polypeptide of SEQ ID NO: 9 herein;
2. (ii) an alpha-amylase comprising an amino acid sequence having at least 60%, at least 70%, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the polypeptide of SEQ ID NO: 9 herein.

[0228] In a preferred embodiment the alpha-amylase is a variant of the alpha-amylase shown in SEQ ID NO: 9 having at least one of the following substitutions or combinations of substitutions: D165M; Y141W; Y141R; K136F; K192R; P224A; P224R; S123H + Y141W; G20S + Y141W; A76G + Y141W; G128D + Y141W; G128D + D143N; P219C + Y141W; N142D + D143N; Y141W + K192R; Y141W + D143N; Y141W + N383R; Y141W + P219C + A265C; Y141W + N142D + D143N; Y141W + K192R V410A; G128D + Y141W + D143N; Y141W + D143N + P219C; Y141W + D143N + K192R; G128D + D143N + K192R; Y141W + D143N + K192R + P219C; G128D + Y141W + D143N + K192R; or G128D + Y141W + D143N + K192R + P219C (using SEQ ID NO: 9 for numbering).

[0229] In an embodiment the alpha-amylase is derived from a *Rhizomucor pusillus* with an *Aspergillus niger* glucoamylase linker and starch-binding domain (SBD), preferably disclosed as SEQ ID NO: 9 herein, preferably having one or more of the following substitutions: G128D, D143N, preferably G128D+D143N (using SEQ ID NO: 9 for numbering), and wherein the alpha-amylase variant present and/or added in saccharification and/or fermentation has at least 75% identity preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, even more preferably at least 93%, most preferably at least 94%, and even most preferably at least 95%, such as even at least 96%, at least 97%, at least 98%, at least 99%, but less than 100% identity to the polypeptide of SEQ ID NO: 9 herein.

[0230] In a preferred embodiment the ratio between glucoamylase and alpha-amylase present and/or added during saccharification and/or fermentation may preferably be in the range from 500:1 to 1:1, such as from 250:1 to 1:1, such as from 100:1 to 1: 1, such as from 100: 2 to 100:50, such as from 100:3 to 100:70.

Starch-Containing Materials

[0231] Any suitable starch-containing starting material may be used in a process of the present invention. The starting material is generally selected based on the desired fermentation product. Examples of starch-containing starting materials, suitable for use in the processes of the present invention, include barley, beans, cassava, cereals, corn, milo, peas, potatoes, rice, rye, sago, sorghum, sweet potatoes, tapioca, wheat, and whole grains, or any mixture thereof. The starch-containing material may also be a waxy or non-waxy type of corn and barley. In a preferred embodiment the starch-containing material is corn. In a preferred embodiment the starch-containing material is wheat.

Fermentation Products

[0232] The term "fermentation product" means a product produced by a method or process including fermenting using a fermenting organism. Fermentation products include alcohols (*e.g.,* ethanol, methanol, butanol); organic acids (*e.g.,* citric acid, acetic acid, itaconic acid, lactic acid, succinic acid, gluconic acid); ketones (*e.g.,* acetone); amino acids (*e.g.,* glutamic acid); gases (*e.g.,* H_2 and CO_2); antibiotics (e.g., penicillin and tetracycline); enzymes; vitamins (*e.g.,* riboflavin, B_{12}, beta-carotene); and hormones. In a preferred embodiment the fermentation product is ethanol, *e.g.,* fuel ethanol; drinking ethanol, *i.e.,* potable neutral spirits; or industrial ethanol or products used in the consumable alcohol industry (*e.g.,* beer and wine), dairy industry (*e.g.,* fermented dairy products), leather industry and tobacco industry. Preferred beer types comprise ales, stouts, porters, lagers, bitters, malt liquors, happoushu, high-alcohol beer, low-alcohol beer, low-calorie beer or light beer. In an preferred embodiment the fermentation product is ethanol.

Fermenting Organisms

[0233] The term "fermenting organism" refers to any organism, including bacterial and fungal organisms, such as yeast and filamentous fungi, suitable for producing a desired fermentation product. Suitable fermenting organisms are able to ferment, *i.e.,* convert, fermentable sugars, such as arabinose, fructose, glucose, maltose, mannose, or xylose, directly or indirectly into the desired fermentation product.

[0234] Examples of fermenting organisms include fungal organisms such as yeast. Preferred yeast include strains of *Saccharomyces,* in particular *Saccharomyces cerevisiae* or *Saccharomyces uvarum;* strains of *Pichia,* in particular *Pichia stipitis* such as *Pichia stipitis* CBS 5773 or *Pichia pastoris;* strains of *Candida,* in particular *Candida arabinofermentans, Candida boidinii, Candida diddensii, Candida shehatae, Candida sonorensis, Candida tropicalis,* or *Candida utilis.* Other fermenting organisms include strains of *Hansenula,* in particular *Hansenula anomala* or *Hansenula polymorpha;* strains of *Kluyveromyces,* in particular

Kluyveromyces fragilis or *Kluyveromyces marxianus*; and strains of *Schizosaccharomyces,* in particular *Schizosaccharomyces pombe.*

[0235] Preferred bacterial fermenting organisms include strains of *Escherichia,* in particular *Escherichia coli,* strains of *Zymomonas,* in particular *Zymomonas mobilis,* strains of *Zymobacter,* in particular *Zymobactor palmae,* strains of *Klebsiella* in particular *Klebsiella oxytoca,* strains of *Leuconostoc,* in particular *Leuconostoc mesenteroides,* strains of *Clostridium,* in particular *Clostridium butyricum,* strains of *Enterobacter,* in particular *Enterobacter aerogenes,* and strains of *Thermoanaerobacter,* in particular *Thermoanaerobacter* BG1 L1 *(*Appl. Microbiol. Biotech. 77: 61-86), *Thermoanarobacter ethanolicus, Thermoanaerobacter mathranii,* or *Thermoanaerobacter thermosaccharolyticum.* Strains of *Lactobacillus* are also envisioned as are strains of *Corynebacterium glutamicum R, Bacillus thermoglucosidaisus,* and *Geobacillus thermoglucosidasius.*

[0236] In an embodiment, the fermenting organism is a C6 sugar fermenting organism, such as a strain of, *e.g., Saccharomyces cerevisiae.*

[0237] In an embodiment, the fermenting organism is a C5 sugar fermenting organism, such as a strain of, *e.g., Saccharomyces cerevisiae.*

[0238] In one embodiment, the fermenting organism is added to the fermentation medium so that the viable fermenting organism, such as yeast, count per mL of fermentation medium is in the range from 10^5 to 10^{12}, preferably from 10^7 to 10^{10}, especially about 5×10^7.

[0239] Yeast is the preferred fermenting organism for ethanol fermentation. Preferred are strains of *Saccharomyces,* especially strains of the species *Saccharomyces cerevisiae,* preferably strains which are resistant towards high levels of ethanol, *i.e.,* up to, *e.g.,* about 10, 12, 15 or 20 vol. % or more ethanol.

[0240] In an embodiment, the C5 utilizing yeast is a Saccharomyces cerevisea strain disclosed in WO 2004/085627.

[0241] In an embodiment, the fermenting organism is a C5 eukaryotic microbial cell concerned in WO 2010/074577 (Nedalco).

[0242] In an embodiment, the fermenting organism is a transformed C5 eukaryotic cell capable of directly isomerize xylose to xylulose disclosed in US 2008/0014620.

[0243] In an embodiment, the fermenting organism is a C5 sugar fermentating cell disclosed in WO 2009/109633.

[0244] Commercially available yeast include LNF SA-1, LNF BG-1, LNF PE-2,and LNF CAT-1 (available from LNF Brazil), RED STAR™ and ETHANOL RED™ yeast (available from Fermentis/Lesaffre, USA), FALI (available from Fleischmann's Yeast, USA), SUPERSTART and

THERMOSACC™ fresh yeast (available from Ethanol Technology, WI, USA), BIOFERM AFT and XR (available from NABC - North American Bioproducts Corporation, GA, USA), GERT STRAND (available from Gert Strand AB, Sweden), and FERMIOL (available from DSM Specialties).

[0245] The fermenting organism capable of producing a desired fermentation product from fermentable sugars is preferably grown under precise conditions at a particular growth rate. When the fermenting organism is introduced into/added to the fermentation medium the inoculated fermenting organism pass through a number of stages. Initially growth does not occur. This period is referred to as the "lag phase" and may be considered a period of adaptation. During the next phase referred to as the "exponential phase" the growth rate gradually increases. After a period of maximum growth the rate ceases and the fermenting organism enters "stationary phase". After a further period of time the fermenting organism enters the "death phase" where the number of viable cells declines.

Fermentation

[0246] The fermentation conditions are determined based on, *e.g.,* the kind of plant material, the available fermentable sugars, the fermenting organism(s) and/or the desired fermentation product. One skilled in the art can easily determine suitable fermentation conditions. The fermentation may be carried out at conventionally used conditions. Preferred fermentation processes are anaerobic processes.

[0247] For example, fermentations may be carried out at temperatures as high as 75°C, e.g., between 40-70°C, such as between 50-60°C. However, bacteria with a significantly lower temperature optimum down to around room temperature (around 20°C) are also known. Examples of suitable fermenting organisms can be found in the "Fermenting Organisms" section above.

[0248] For ethanol production using yeast, the fermentation may go on for 24 to 96 hours, in particular for 35 to 60 hours. In an embodiment the fermentation is carried out at a temperature between 20 to 40°C, preferably 26 to 34°C, in particular around 32°C. In an embodiment the pH is from pH 3 to 6, preferably around pH 4 to 5.

[0249] Other fermentation products may be fermented at temperatures known to the skilled person in the art to be suitable for the fermenting organism in question.

[0250] Fermentation is typically carried out at a pH in the range between 3 and 7, preferably from pH 3.5 to 6, such as around pH 5. Fermentations are typically ongoing for 6-96 hours.

[0251] The processes of the invention may be performed as a batch or as a continuous process. Fermentations may be conducted in an ultrafiltration system wherein the retentate is

held under recirculation in the presence of solids, water, and the fermenting organism, and wherein the permeate is the desired fermentation product containing liquid. Equally contemplated are methods/processes conducted in continuous membrane reactors with ultrafiltration membranes and where the retentate is held under recirculation in presence of solids, water, and the fermenting organism(s) and where the permeate is the fermentation product containing liquid. After fermentation the fermenting organism may be separated from the fermented slurry and recycled.

Fermentation Medium

[0252] The phrase "fermentation media" or "fermentation medium" refers to the environment in which fermentation is carried out and comprises the fermentation substrate, that is, the carbohydrate source that is metabolized by the fermenting organism(s).

[0253] The fermentation medium may comprise other nutrients and growth stimulator(s) for the fermenting organism(s). Nutrient and growth stimulators are widely used in the art of fermentation and include nitrogen sources, such as ammonia; vitamins and minerals, or combinations thereof.

Recovery

[0254] Subsequent to fermentation, the fermentation product may be separated from the fermentation medium. The fermentation medium may be distilled to extract the desired fermentation product or the desired fermentation product may be extracted from the fermentation medium by micro or membrane filtration techniques. Alternatively, the fermentation product may be recovered by stripping. Methods for recovery are well known in the art.

[0255] The present invention is further described by the following examples.

EXAMPLES

Enzymes

[0256] Enzymes for DNA manipulations (e.g. restriction endonucleases, ligases etc.) were obtained from New England Biolabs, Inc. and were used according to the manufacturer's instructions.

Media and reagents

[0257] The following media and reagents were used unless otherwise specified:
Chemicals used for buffers and substrates were commercial products of analytical grade. Cove: 342.3 g/L Sucrose, 20 ml/L COVE salt solution, 10mM Acetamide, 30 g/L noble agar.

[0258] Cove top agar: 342.3 g/L Sucrose, 20 ml/L COVE salt solution, 10mM Acetamide, 10 g/L low melt agarose. Cove-N plates are composed of 30 g sucrose, 20 ml Cove salt solution, 3g $NaNO_3$, and 30 g noble agar and water to 1 litre. COVE salt solution are composed of 26 g KCl, 26 g $MgSO_4$ $7H_2O$, 76 g KH_2PO_4 and 50ml Cove trace metals and water to 1 litre. Trace metal solution for COVE are composed of 0.04 g NaB_4O_7 $10H_2O$, 0.4 g $CuSO_4$ $5H_2O$, 1.2 g $FeSO_4 7H_2O$, 1.0 g $MnSO_4$ H_2O, 0.8 g Neutral amylase II $MoO_2 2H_2O$, and 10.0 g $ZnSO_4$ $7H_2O$ and water to 1 litre. ¼ YPM composed of 2.5 g yeast extract, 5 g pepton and 5 g maltose (pH 4.5) and water to 1 litre. STC buffer was composed of 0.8 M sorbitol, 25 mM Tris (pH 8), and 25 mM $CaCl_2$ and water to 1litre. STPC buffer composed of 40% PEG4000 in STC buffer. MLC composed of 40 g Glucose, 50 g Soybean powder, 4 g/ Citric acid (pH 5.0) and water to 1 litre.

Purchased material (E.coli, plasmid and kits)

[0259] *E.coli* DH5-alpha (Toyobo) was used for plasmid construction and amplification. Amplified plasmids were recovered with Qiagen Plasmid Kit (Qiagen). QIAquickTM Gel Extraction Kit (Qiagen) was used for the purification of PCR fragments and extraction of DNA fragment from agarose gel.

Strains

[0260] The expression host strain *Aspergillus* niger described is a derivative of NN059203. NN059203 was isolated by Novozymes and described in WO12160093 and is a derivative of *Aspergillus niger* NN049184 which was isolated from soil.

Transformation of Aspergillus

[0261] Transformation of *Aspergillus* species can be achieved using the general methods for yeast transformation. The preferred procedure for the invention is described below. The *Aspergillus niger host* strain was inoculated into 100 ml YPG medium supplemented with 10 mM uridine and incubated for 16 hrs at 32°C at 80 rpm. Pellets were collected and washed with 0.6 M KCl, and resuspended in 20 ml 0.6 M KCl containing a commercial glucanase product (GLUCANEX™, Novozymes A/S, Bagsværd, Denmark) at a final concentration of 20 mg per ml. The suspension was incubated at 32°C with shaking (80 rpm) until protoplasts were

formed, and then washed twice with STC buffer. The protoplasts were counted with a hematometer and resuspended and adjusted in an 8:2:0.1 solution of STC:STPC:DMSO to a final concentration of 2.5×10^7 protoplasts/ml. Approximately 4pg of plasmid DNA was added to 100 pl of the protoplast suspension, mixed gently, and incubated on ice for 30 minutes. One ml of SPTC was added and the protoplast suspension was incubated for 20 minutes at 37°C. After the addition of 10 ml of 50°C Cove top agarose, the reaction was poured onto Cove agar plates and the plates were incubated at 32°C for 5 days.

PCR amplification

[0262]

PrimeSTAR® HS (Premix) 10 µl

Template DNA (50-100 ng/ pl) 1 µl

Forward primer (100pM) 1 µl

Reverse primer (100 pM) 1 µl

Distilled water to 20 µl

PCR conditions

[0263]

1. 94°C 2 min

2. 94°C 10 sec

3. 57°C 5 sec

4. 72°C 20 sec

Repeat 2-4, 30 cycles

5. 72°C 30 sec

MTP cultivation for enzyme production

[0264] Spores of Aspergillus libraries were inoculated in 0.5-1ml of 1/4YPM media in 96 deep well plate and cultivated at 30°C for 2-3 days at 600rpm.

Enzyme assay

Zein plate assay

[0265] Culture supernatants were applied on 0.05-0.1% of zein (Sigma) plate (20mM sodium acetate buffer, pH4.5) and incubated at appropriate temperatures (30-60 degree C).

Suc-AAPF-pna analysis

[0266] Culture supernatants pre-incubated at appropriate temperatures (50 to 60 degree C and 4 degree C as a control) are measured for protease activity by AAPF assay using N-Succinyl-Ala-Ala-Pro-Phe p-nitroanilide (SIGMA-ALDRICH).

Assay:

[0267]

1. 1) Add 25µl samples to wells of 384 microtiterplate (MTP)
2. 2) Add 25µl pNA substrate working soln. to 384 MTP
3. 3) Read 405nm (zero point)
4. 4) Incubate 37°C, 60min (if the color is not developed well, continue incubation)
5. 5) Read 405nm (zero point)

DNA isolation from Aspergillus clones

[0268] Inserted DNAs of Aspergillus strains were recovered by direct PCR method described below or PCR on the isolated DNA by chromosomal DNA purification kit (FastDNA SPIN Kit for Soil, MP biomedicals, #6560-200) using a primer pair, insert rescue F and R.

insert rescue F (SEQ ID NO: 11)AATCTCAGAACACCAATATC

insert rescue R (SEQ ID NO: 12) AACACTATGCGTTATCGTAC

[0269] The amplified DNAs were purified by agarose gel electrophoresis followed to QIAquick Gel Extraction kit (Qiagen) for sequencing analysis to check the quality of constructed libraries. Colony PCR was carried out as following:
Conidias from strains were transferred to a 1.5 ml tube and 500µl of TE-buffer was added and mixed briefly.

[0270] It was diluted 10-20 times in water and one µl of the dilution was used as template for PCR.

Purification

[0271] Purification of the variants of *Meripilus giganteus* serine protease belonging to family 53 herein denoted as MgProtIII variants was carried out by two steps, desalting column and cation exchange chromatography column. Finally, the sample was buffer exchanged and concentrated in 20 mM succinate buffer pH 4.0 using a 30 kDa centrifugal concentrator (Sartorius AG).

TSA (Thermal Shift Assay)

[0272] Purified enzyme was diluted with 50 mM sodium acetate buffer pH 4.5 to 0.75 mg/ml and 10 µl of that were mixed with 15 µl of SYPRO Orange (Invitrogen) diluted with Milli-Q water and 5 µl of 30 mM bacitracin solution dissolved in 50 mM sodium acetate buffer pH 4.5. Thirty microliters of mixture solution was transfer to LightCycler 480 Multiwell Plate 96 (Roche Diagnostics) and the plate was sealed.

Equipment parameters of TSA:

[0273]

Apparatus: LightCycler 480 Real-Time PCR System (Roche Applied Science)

Scan rate: 0.02°C/sec

Scan range: 37 - 96°C

Integration time: 1.0 sec

Excitation wave length 465 nm

Emission wave length 580 nm

[0274] The obtained fluorescence signal was normalized into a range of 0 and 1. The Thermal denaturation temperature, Td, was defined as the temperature where the normalized value is closest to 0.5.

EXAMPLE 1: Library construction

Plasmid library construction using In-fusion cloning (Clontech)

[0275] An expression vector, pFLP-MgProIII disclosed in WO1260093 Figure 5, which contains the target protease gene (shown as SEQ 1) instead of glucoamylase gene and amd S marker gene instead of pyr G marker gene, was digested with appropriate restriction enzymes (XhoI BsiW1 for pFRT-GIAMG) to cut out the protease gene.

[0276] Two PCRs were carried out for a library construction using 2 primer pairs, a forward degenerated primer and a primer having more than 15bp overlapping with an expression vector (vector F described below), and vector F primer and a reverse primer having 15bp overlapping with the degenerated primer using the expression vector as a template.

Vector R 25mer (SEQ ID NO: 13) TAAGTGGAGGGAAAAACACTATGCG

Vector F 32mer (SEQ ID NO: 14) GCTTGGAGCAACAATCTCAGAACACCAATATC

[0277] One of the examples of primers for a library is shown below:

F111X F 27mer (SEQ ID NO: 15) ATCTCCGTCGGCGACGACNNKCAGGAT

F111X R 20mer (SEQ ID NO: 16) GTCGCCGACGGAGATGAACG

[0278] The digested vector and PCR fragments were mixed with In-Fusion mix and transformed into *E.coli* DH5alpha. Obtained *E.coli* transformants were pooled and plasmids were extracted to use for *Aspergillus* library construction.

Aspergillus transformation to construct a library in Aspergillus

[0279] One µg of each plasmid library was transformed into *A. niger* host strain. *Aspergillus* transformants were isolated in a 96 well-MTP containing COVE-N gly agar (100ul/well), cultivate at 32°C for 1 week to have enough sporulation. 100µl/well of 0.01% tween 20 was

added to the each well, suspended with spores and the suspension was inoculated in a 96 well-MTP containing YPG and cultivated for 3 days at 30°C with shaking to have Aspergillus culture library. They were used for further library screening works.

Library screening

[0280] Constructed Aspergillus libraries were cultivated in 96 well MTP and the culture supernatants were spotted on zein plates at appropriate temepratures. Positive variants were tested by Suc-AAPF-pna analysis and variants having higher residual activities were identified. Positive variants were cultivated in shake flasks and samples were used for further purification and characterization.

DNA isolation from Aspergillus clones

[0281] Inserted DNAs of Aspergillus strains were recovered by direct PCR method described below or PCR on the isolated DNA by chromosomal DNA purification kit (FastDNA SPIN Kit for Soil, MP biomedicals, #6560-200) using a primer pair, insert rescue F and R.

insert rescue F (SEQ ID NO: 11)AATCTCAGAACACCAATATC

insert rescue R (SEQ ID NO: 12) AACACTATGCGTTATCGTAC

[0282] The amplified DNAs were purified by agarose gel electrophoresis followed to QIAquick Gel Extraction kit (Qiagen) for sequencing analysis to check the quality of constructed libraries.

RESULTS

[0283] Table 1 lists the positive variants identified by Suc-AAPF-pna analysis. Samples were incubated at certain temperatures for 30 minutes and their remaining activities were measured by AAPF assay. The residual activities in tables below are described as relative activity to ones incubated at 4 degree C.

JMgP ID	Modification	<Residual activity (%)>	
		57°C, 30min	58°C, 30min
WT	-	8	8
JMgP006	N115L	10	9
JMgP019	N115L S183L S187L	18	15
JMgP033	N115L Q182G	36	4

JMgP ID	Modification	<Residual activity (%)>	
		57°C, 30min	58°C, 30min
JMgP076	N115D	45	4
JMgP071	N115L S183L S187L P348A	22	11
JMgP ID	-	55°C, 30min	56°C, 30min
WT	-	19	4
JMgP009	S183P	31	10
JMgP033	N115L Q182G	38	17
JMgP034	N115L Q182R	60	35
JMgP058	N115L S183L S187L V209W S210*	47	19
JMgP059	N115L S183L S187L V209L S210*	39	20
JMgP064	N115L S183L S187L S317G S318*	51	37
JMgP065	N115L S183L S187L S317S S318*	47	22
JMgP066	N115L S183L S187L S317A S318*	47	33
JMgP075	S183L V209L S210*	49	41
JMgP ID		55°C, 30min	56°C, 30min
JMgP009	S183P	39.4	19
lib19-4	S183P E74W	32.1	13.7
lib22-2	S183P E81A	46.1	23.5
JMgP083	E81R S183P	70.9	53.5
JMgP084	E81K S183P	73.5	55.6
lib22-11	S183P E81E	36.8	15.3
JMgP ID		55°C, 30min	56°C, 30min
JMgP009	S183P	46.8	24.7
JMgP094	S183P Q154V	67.4	49.4
JMgP095	S183P Q142W	66.1	50.6
JMgP096	Q142R S183P	99	100
JMgP097	S183P T146A	65.8	47.4
JMgP098	S183P T146W	83.8	78.1
JMgP099	S183P I228R	67	49.5
JMgP100	S183P D267N	81.9	76.6
JMgP101	S183P S272V	81.8	77.5
JMgP103	S183P S272R	88.6	83.2

JMgP ID	Modification	<Residual activity (%)>	
		57°C, 30min	58°C, 30min
JMgP120	T146Y S183P	64	52
JMgP ID		55°C, 30min	56°C, 30min
JMgP009	S183P	48	27
JMgP030	D280N	62	49
JMgP087	D109P S183P V209L S210*	63	49
JMgP089	E81R S183P V209L S210*	69	62
JMgP091	D109P V209L S210*	47	28
JMgP092	N115D V209L S210*	54	36
JMgP093	E81R V209L S210*	56	42
JMgP104	V209L S210* S317A S318*	58	45
JMgP106	E81R V209L S210* S317A S318*	65	55
JMgP115	T146W D280N	60	50
JMgP118	T146W S183P D280N	73	29
JMgP120	T146Y S183P	81	37
JMgP134	S183P S294A	85	58
JMgP137	S183P T362A	53	32
JMgP140	S183P S294A	73	56
		55°C, 30min	56°C, 30min
JMgP009	S183P	50	29
JMgP088	N115D S183P V209L S210*	41	34
JMgP096	Q142R S183P	73	71
JMgP123	S183P E212E	48	32
JMgP127	I84C S183P S272C	71	67
		57°C, 30min	58°C, 30min
JMgP096	Q142R S183P	78	76
JMgP229	Q142R N145G T146E S183P	86	87
JMgP230	Q142R N145Q T146D S183P	92	92
JMgP231	Q142R N145V T146E S183P	92	91
JMgP232	Q142R N145D T146E S183P	89	90
JMgP233	Q142R N145K T146E S183P	92	94

JMgP ID	Modification	<Residual activity (%)>	
		57°C, 30min	58°C, 30min
JMgP234	Q142R N145A T146D S183P	84	82
JMgP236	I39M Q142R S183P	88	92
JMgP235	Q142R N145E T146E S183P	90	87
JMgP237	I39R Q142R S183P	82	82
JMgP238	I39L Q142R S183P	91	98
		60°C, 30min	62°C, 30min
JMgP096	Q142R S183P	84	45
JMgP245	I39C Q142R S183P	125	105
		57°C, 30min	60°C, 30min
JMgP252	E117D Q142R S183P	90	61
JMgP096	Q142R S183P	88	43
		57°C, 30min	60°C, 30min
lib81-1	S60D Q142R S183P	96	99
JMgP096	Q142R S183P	96	60

EXAMPLE 2: Purification and Thermal shift assay (TSA) analysis

Purification

[0284] Purification of MgProtIII variants was carried out by two steps, desalting column and cation exchange chromatography column. Finally, the sample was buffer exchanged and concentrated in 20 mM succinate buffer pH 4.0 using a 30 kDa centrifugal concentrator (Sartorius AG).

TSA

[0285] Purified enzyme was diluted with 50 mM sodium acetate buffer pH 4.5 to 0.75 mg/ml and 10 µl of that were mixed with 15 µl of SYPRO Orange (Invitrogen) diluted with Milli-Q water and 5 µl of 30 mM bacitracin solution dissolved in 50 mM sodium acetate buffer pH 4.5. Thirty microliters of mixture solution was transfer to LightCycler 480 Multiwell Plate 96 (Roche

Diagnostics) and the plate was sealed.

Equipment parameters of TSA:

[0286]

Apparatus: LightCycler 480 Real-Time PCR System (Roche Applied Science)

Scan rate: 0.02°C/sec

Scan range: 37 - 96°C

Integration time: 1.0 sec

Excitation wave length 465 nm

Emission wave length 580 nm

[0287] The obtained fluorescence signal was normalized into a range of 0 and 1. The Td was defined as the temperature where the normalized value is closest to 0.5.

[0288] Result: The TSA data are listed in TABLE 2.

Sample	Modification	Td [°C]
MgProtIII (wt)	-	58.67
JMgP009	S183P	60.28
JMgP030	D280N	61.15
JMgP081	K57R S183P	59.67
JMgP087	D109P S183P V209L S210*	62.37
JMgP089	E81R S183P V209L S210*	63.03
JMgP093	E81R V209L S210*	62.01
JMgP094	Q154V S183P	61.06
JMgP095	Q142W S183P	61.39
JMgP096	Q142R S183P	65.77
JMgP097	T146A S183P	61.18
JMgP098	T146W S183P	62.34
JMgP099	S183P I228R	60.59
JMgP100	S183P D267N	62.43
JMgP101	S183P S272V	61.06
JMgP106	E81R V209L S210* S317A S318*	62.15

Sample	Modification	Td [°C]
JMgP108	S183P T328C K343C	60.59
JMgP110	S183P G320C A363C	59.71
JMgP115	T146W D280N	63.33
JMgP116	T146W S183L D280N	63.27
JMgP117	T146W	62.46
JMgP118	T146W S183P D280N	63.72
JMgP120	T146Y S183P	62.13
JMgP122	S183P Q207R	61.36
JMgP126	S50C S183P V271C	66.92
JMgP127	I84C S183P S272C	64.44
JMgP130	Q142W T146W S183P	61.94
JMgP132	Q142W T146W S183P D280N	63.00
JMgP134	S183P S294A	61.61
JMgP136	S183P K321G	60.63
JMgP137	S183P T362A	60.87
JMgP141	Q182G	59.88
JMgP144	Q142W T146W Q182R	62.13
JMgP147	S272V	63.18
JMgP148	S272R	61.67
JMgP157	S60P	61.74
JMgP167	D109N D110N	60.64
JMgP173	F111P	59.43
JMgP175	G128A	59.95
JMgP203	G278S	59.95
JMgP206	S318N K321A A322S	60.23
JMgP214	E81R T146W	64.02
JMgP215	E81R Q142R S183P	66.08
JMgP216	E81R Q142W S183P	62.81
JMgP218	S183P G274G	61.03
JMgP220	E81R	62.34
JMgP223	N124L Q142R S183P	65.69
JMgP224	N124W Q142R S183P	64.09
JMgP225	N124Q Q142R S183P	65.38
JMgP231	Q142R N145V T146E S183P	65.79

Sample	Modification	Td [°C]
JMgP232	Q142R N145D T146E S183P	65.37
JMgP234	Q142R N145A T146D S183P	65.52

REFERENCES CITED IN THE DESCRIPTION

Cited references

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US5231017A [0003]
- WO2003066826A [0004]
- WO2007145912A [0005]
- WO2014037438A [0006]
- US62232903 [0007]
- US20040171154 [0094]
- WO9517413A [0097]
- WO9522625A [0097]
- US5223409A [0097]
- WO9206204A [0097]
- WO9600787A [0104] [0136]
- WO0056900A [0104] [0104] [0104]
- WO9533836A [0114]
- WO0024883A [0125] [0125]
- EP238023A [0136]
- WO2011066576A [0150] [0150] [0216] [0216]
- WO2011068803A [0152] [0152] [0218] [0218]
- WO2012064351A [0155] [0221]
- WO2013006756A [0160] [0226]
- WO99019467A [0177]
- WO9919467A [0183]
- US4335208A [0186]

- US3912590A [0192]
- EP252730A [0192]
- EP063909A [0192]
- US4316956A [0198]
- WO2004085627A [0240]
- WO2010074577A [0241]
- US20080014620A [0242]
- WO2009109633A [0243]
- WO12160093A [0260]
- WO1260093A [0275]

Non-patent literature cited in the description

- **ANSON, M.L.MIRSKY, A.E.**J. Gen. Physiol., 1932, vol. 16, 59- [0015]
- **ANSON, M.L.**J. Gen. Physiol., 1938, vol. 22, 79- [0015]
- Biochem. J, 1993, vol. 290, 205-218 [0018]
- **RAWLINGS, N.D.BARRETT, A.J.BATEMAN, A.**MEROPS: the peptidase databaseNucl. Acids Res., 2010, vol. 38, D227-D233 [0018]
- **NEEDLEMANWUNSCH**J. Mol. Biol., 1970, vol. 48, 443-453 [0042] [0049]
- **RICE et al.**EMBOSS: The European Molecular Biology Open Software SuiteTrends Genet., 2000, vol. 16, 276-277 [0042] [0049]
- **EDGAR**Nucleic Acids Research, 2004, vol. 32, 1792-1797 [0050]
- **KATOHKUMA**Nucleic Acids Research, 2002, vol. 30, 3059-3066 [0050]
- **KATOH et al.**Nucleic Acids Research, 2005, vol. 33, 511-518 [0050]
- **KATOHTOH**Bioinformatics, 2007, vol. 23, 372-374 [0050]
- **KATOH et al.**Methods in Molecular Biology, 2009, vol. 537, 39-64 [0050]
- **KATOHTOH**Bioinformatics, 2010, vol. 26, 1899-1900 [0050]
- **THOMPSON et al.**Nucleic Acids Research, 1994, vol. 22, 4673-4680 [0050]
- **LINDAHLELOFSSON**J. Mol. Biol., 2000, vol. 295, 613-615 [0051]
- **ATSCHUL et al.**Nucleic Acids Res., 1997, vol. 25, 3389-3402 [0051]
- **JONES**J. Mol. Biol., 1999, vol. 287, 797-815 [0051]
- **MCGUFFINJONES**Bioinformatics, 2003, vol. 19, 874-881 [0051]
- **GOUGH et al.**J. Mol. Biol., 2000, vol. 313, 903-919 [0051]
- **HOLMSANDER**Proteins, 1998, vol. 33, 88-96 [0052]
- **SHINDYALOVBOURNE**Protein Engineering, 1998, vol. 11, 739-747 [0052]
- **HOLMPARK**Bioinformatics, 2000, vol. 16, 566-567 [0052]
- **H. NEURATHR.L. HILL**The ProteinsAcademic Press19790000 [0073]
- **CUNNINGHAMWELLS**Science, 1989, vol. 244, 1081-1085 [0074]
- **HILTON et al.**J. Biol. Chem., 1996, vol. 271, 4699-4708 [0074]
- **DE VOS et al.**Science, 1992, vol. 255, 306-312 [0074]

- **SMITH et al.**J. Mol. Biol., 1992, vol. 224, 899-904 [0074]
- **WLODAVER et al.**FEBS Lett., 1992, vol. 309, 59-64 [0074]
- **SAMBROOK et al.**Molecular Cloning, A Laboratory ManualCold Spring Harbor19890000 [0081]
- **COOPER et al.**EMBO J., 1993, vol. 12, 2575-2583 [0088]
- **DAWSON et al.**Science, vol. 266, 776-779 [0088]
- **MARTIN et al.**J. Ind. Microbiol. Biotechnol., 2003, vol. 3, 568-576 [0089]
- **SVETINA et al.**J. Biotechnol., 2000, vol. 76, 245-251 [0089]
- **RASMUSSEN-WILSON et al.**Appl. Environ. Microbiol., 1997, vol. 63, 3488-3493 [0089]
- **WARD et al.**Biotechnology, 1995, vol. 13, 498-503 [0089]
- **CONTRERAS et al.**Biotechnology, 1991, vol. 9, 378-381 [0089]
- **EATON et al.**Biochemistry, 1986, vol. 25, 505-512 [0089]
- **COLLINS-RACIE et al.**Biotechnology, 1995, vol. 13, 982-987 [0089]
- **CARTER et al.**Proteins: Structure, Function, and Genetics, 1989, vol. 6, 240-248 [0089]
- **STEVENS**Drug Discovery World, 2003, vol. 4, 35-48 [0089]
- **SCHERERDAVIS**Proc. Natl. Acad. Sci. USA, 1979, vol. 76, 4949-4955 [0093]
- **BARTON et al.**Nucleic Acids Res., 1990, vol. 18, 7349-4966 [0093]
- **STORICI et al.**Nature Biotechnol., 2001, vol. 19, 773-776 [0094]
- **KREN et al.**Nat. Med., 1998, vol. 4, 285-290 [0094]
- **CALISSANOMACINO**Fungal Genet. Newslett., 1996, vol. 43, 15-16 [0094]
- **TIAN et al.**Nature, 2004, vol. 432, 1050-1054 [0096]
- **REIDHAAR-OLSONSAUER**Science, 1988, vol. 241, 53-57 [0097]
- **BOWIESAUER**Proc. Natl. Acad. Sci. USA, 1989, vol. 86, 2152-2156 [0097]
- **LOWMAN et al.**Biochemistry, 1991, vol. 30, 10832-10837 [0097]
- **DERBYSHIRE et al.**Gene, 1986, vol. 46, 145- [0097]
- **NER et al.**DNA, 1988, vol. 7, 127- [0097]
- **NESS et al.**Nature Biotechnology, 1999, vol. 17, 893-896 [0098]
- **GEMS et al.**Gene, 1991, vol. 98, 61-67 [0125]
- **CULLEN et al.**Nucleic Acids Res., 1987, vol. 15, 9163-9175 [0125]
- **HAWKSWORTH et al.**Ainsworth and Bisby's Dictionary of The FungiCAB International, University Press19950000 [0130]
- Biology and Activities of YeastSoc. App. Bacteriol. Symposium Series19800000 [0131]
- **YELTON et al.**Proc. Natl. Acad. Sci. USA, 1984, vol. 81, 1470-1474 [0136]
- **CHRISTENSEN et al.**BiolTechnology, 1988, vol. 6, 1419-1422 [0136]
- **MALARDIER et al.**Gene, 1989, vol. 78, 147-156 [0136]
- Guide to Yeast Genetics and Molecular Biology**BECKERGUARENTE**Methods in EnzymologyAcademic Press, Inc.vol. 194, 182-187 [0136]
- **ITO et al.**J. Bacteriol., 1983, vol. 153, 163 [0136]
- **HINNEN et al.**Proc. Natl. Acad. Sci. USA, 1978, vol. 75, 1920- [0136]
- Protein PurificationVCH Publishers19890000 [0141]
- **GORINSTEINLII**Starch/Stärke, 1992, vol. 44, 12461-466 [0205]
- Appl. Microbiol. Biotech., vol. 77, 61-86 [0235]

KRAV

1. Proteasevariant omfattende en modifikation ved en eller flere positioner svarende til positioner 39, 50, 57, 60, 74, 81, 84, 109, 110, 111, 115, 117, 124, 142, 145, 146, 154, 182, 183, 187, 207, 209, 210, 228, 267, 271, 272, 274, 280, 294, 317, 318, 320, 321, 322, 328, 343, 348, 362 eller 363 i polypeptidet ifølge SEQ ID NO: 3, hvor varianten har proteaseaktivitet, og hvor varianten har mindst 90%, mindst 95%, mindst 96%, mindst 97%, mindst 98% eller mindst 99% men mindre end 100% sekvensidentitet med det mature polypeptid ifølge SEQ ID NO: 3, og hvor varianten har en forøget termostabilitet sammenlignet med proteasen ifølge SEQ ID NO: 3, hvor den forøgede termostabilitet er forøget restaktivitet målt efter inkubation i 30 min. ved en temperatur i området fra 55 til 60 grader Celsius, og hvor varianten omfatter mindst én af de følgende modifikationer eller kombination af modifikationer:

N115L;
S183P;
N115L + Q182G;
N115L + Q182R;
E81R + S183P;
E81K + S183P;
S183P + Q154V;
S183P + Q142W;
Q142R + S183P;
S183P + T146A;
S183P + T146W;
S183P + I228R;
S183P + D267N;
S183P + S272V;
S183P + S272R;
T146Y + S183P;
S183P + S294A;
S183P + T362A;
S183P + S294A;
S183P + E74W;
S183P + E81A;
N115L + S183L + S187L;
S183L + V209L + S210 deletion;
D109P + V209L + S210 deletion;
N115D + V209L + S210 deletion;
E81R + V209L + S210 deletion;
D109P + V209L + S210 deletion;

N115D + V209L + S210 deletion;
E81R + V209L + S210 deletion;
T146W + S183P + D280N;
I84C + S183P + S272C;
I39M + Q142R + S183P;
I39R + Q142R + S183P;
I39L + Q142R + S183P;
I39C + Q142R + S183P;
E117D + Q142R + S183P;
S60D + Q142R + S183P;
N115L + S183L +S187L + P348A;
D109P + S183P + V209L + S210 deletion;
N115D + S183P + V209L + S210 deletion;
E81R + S183P + V209L +S210 deletion;
V209L + S210 deletion + S317A + S318 deletion;
Q142R + N145G + T146E + S183P;
Q142R + N145Q + T146D + S183P;
Q142R + N145V + T146E + S183P;
Q142R + N145D + T146E + S183P;
Q142R + N145K + T146E + S183P;
Q142R + N145A + T146D +S183P;
Q142R + N145E + T146E + S183P;
N115L + S183L + S187L + V209W + S210 deletion;
N115L + S183L + S187L + V209L + S210 deletion;
N115L + S183L + S187L + S317G + S318 deletion;
N115L + S183L + S187L + S317S + S318 deletion;
N115L +S183L + S187L + S317A + S318 deletion;
E81R + V209L + S210 deletion + S317A + S318 deletion.

2. Variant ifølge krav 1, hvor varianten har en restaktivitet på mindst 10%, fortrinsvis mindst 12%, mere fortrinsvis mindst 15%, målt efter inkubation i 30 min. ved 56°C.

3. Proteasevariant omfattende en modifikation ved en position svarende til position 50, 57, 81, 84, 109, 110, 111, 124, 142, 145, 146, 154, 182, 183, 207, 209, 210, 228, 267, 271, 272, 274, 280, 294, 317, 318, 320, 321, 322, 328, 343, 362 eller 363 i polypeptidet ifølge SEQ ID NO: 3, hvor varianten har proteaseaktivititet, og hvor varianten har mindst 90%, mindst 95%, mindst 96%, mindst 97%, mindst 98% eller mindst 99% men mindre end 100% sekvensidentitet med polypeptidet ifølge SEQ ID NO: 3, hvor

varianten har forøget termostabilitet sammenlignet med proteasen ifølge SEQ ID NO: 3, hvor varianten omfatter mindst én af de følgende modifikationer eller kombination af modifikationer:

S183P;
K57R +S183P;
D109P +S183P+ V209L +S210 deletion;
E81R+ S183P +V209L +S210 deletion;
E81R +V209L +S210 deletion;
Q154V +S183P;
Q142W +S183P;
Q142R +S183P;
T146A +S183P;
T146W +S183P;
S183P +I228R;
S183P +D267N;
S183P +S272V;
E81R +V209L +S210 deletion +S317A +S318 deletion;
S183P +T328C +K343C;
S183P +G320C+ A363C;
T146W +S183L D+280N;
T146W +S183P +D280N;
T146Y +S183P;
S183P +Q207R;
S50C +S183P +V271C;
I84C +S183P+ S272C;
Q142W +T146W +S183P;
Q142W +T146W +S183P+ D280N;
S183P +S294A;
S183P +K321G;
S183P +T362A;
Q182G;
Q142W +T146W +Q182R;
S272V;
D109N +D110N;
F111P;
S318N +K321A +A322S;
E81R+ T146W;
E81R +Q142R +S183P;

E81R +Q142W +S183P;
S183P +G274G;
E81R;
N124L +Q142R +S183P;
N124W +Q142R +S183P;
N124Q +Q142R +S183P;
Q142R + N145V + T146E +S183P;
Q142R +N145D +T146E + S183P;
Q142R + N145A +T146D +S183P; og hvor den forøgede termostabilitet, målt som Td ved TSA-assay, er mindst 59°C.

4. Polynukleotid, der koder for varianten ifølge ethvert af kravene 1-3.

5. Nukleinsyrekonstrukt, der omfatter polynukleotidet ifølge krav 4.

6. Ekspressionsvektor, der omfatter polynukleotidet ifølge krav 4.

7. Rekombinant værtscelle, der omfatter polynukleotidet ifølge krav 4.

8. Fremgangsmåde til fremstilling af en proteasevariant ifølge ethvert af kravene 1-3, hvilken fremgangsmåde omfatter dyrkning af værtscellen ifølge krav 7 under betingelser, der er egnede til ekspression af varianten; og, valgfrit, indvinding af varianten.

9. Sammensætning, der omfatter varianten ifølge ethvert af kravene 1-3.

10. Sammensætning ifølge krav 9, hvilken sammensætning yderligere omfatter en glucoamylase og, valgfrit, en svampe-alfa-amylase.

11. Fremgangsmåde til fremstilling af et fermenteringsprodukt fra stivelsesholdigt materiale, hvilken fremgangsmåde omfatter samtidig forsukring og fermentering af stivelsesholdigt materiale under anvendelse af en kulhydratkilde, der danner enzymer, og en fermenteringsorganisme ved en temperatur under den indledende gelatineringstemperatur for det stivelsesholdige materiale i nærværelse af en proteasevariant ifølge ethvert af kravene 1-3.

12. Fremgangsmåde til fremstilling af et fermenteringsprodukt fra stivelsesholdigt materiale omfattende trinnene:

a) flydendegørelse af stivelsesholdigt materiale i nærværelse af en alfa-amylase;

b) forsukring af det flydendegjorte materiale opnået i trin (a) under anvendelse af en glucoamylase;

c) fermentering under anvendelse af en fermenteringsorganisme; hvor proteasevarianten ifølge ethvert af kravene 1-3 er til stede under trin b) og/eller c).

13. Fremgangsmåde ifølge ethvert af kravene 11-12, hvor fermenteringsproduktet er ethanol, og fermenteringsorganismen er Saccharomyces cerevisiae.

14. Værtscelle ifølge krav 7, hvilken værtscelle udtrykker varianterne ifølge ethvert af kravene 1-3, hvor værtscellen er en gærcelle, fortrinsvis en Saccharomyces, såsom *Saccharomyces cerevisiae.*

15. Fremgangsmåde ifølge ethvert af kravene 11-12, hvor værtscellen ifølge 14, anvendes som fermenteringsorganisme i fermenteringstrinnet, og fermenteringsproduktet er ethanol.

SEKVENSLISTE

Sekvenslisten er udeladt af skriftet og kan hentes fra det Europæiske Patent Register.

The Sequence Listing was omitted from the document and can be downloaded from the European Patent Register.