



- (51) International Patent Classification:
A23K 1/165 (2006.01)
- (21) International Application Number:
PCT/EP2014/052218
- (22) International Filing Date:
5 February 2014 (05.02.2014)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
13154179.9 6 February 2013 (06.02.2013) EP
- (71) Applicant: NOVOZYMES A/S [DK/DK]; Krogshoejvej 36, DK-2880 Bagsvaerd (DK).
- (72) Inventors: OLINSKI, Robert Piotr; Bymidten 69, 2. th., DK-3500 Vaerloese (DK). OESTERGAARD, Peter Rahbek; Kvaedevej 111, DK-2830 Virum (DK). PONTOPIDAN, Katrine; Uggeloese Bygade 92A, DK- Lyngø (DK). HOFF, Tine; Soelleroedgaardsvej 38, DK-2840 Holte (DK).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,

HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))



(54) Title: POLYPEPTIDES HAVING PROTEASE ACTIVITY

(57) Abstract: The present invention relates to isolated polypeptides having protease activity, and the use of isolated polypeptides having protease activity in animal feed. It also relates to the use of isolated nucleic acid sequences encoding the proteases in the recombinant production of isolated polypeptides having protease activity and isolated nucleic acid sequences encoding the proteases. The invention also relates to nucleic acid constructs, vectors, and host cells, including plant and animal cells, comprising the nucleic acid sequences, as well as methods for producing and using the proteases, particularly using the proteases in animal feed.

POLYPEPTIDES HAVING PROTEASE ACTIVITY

Reference to a Sequence Listing

This application contains a Sequence Listing in computer readable form, which is incorporated herein by reference.

5

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates to isolated polypeptides having protease activity, and the use of isolated polypeptides having protease activity in animal feed. It also relates to the use of isolated nucleic acid sequences encoding the proteases in the recombinant production of isolated polypeptides having protease activity and isolated nucleic acid sequences encoding the proteases. The invention also relates to nucleic acid constructs, vectors, and host cells, including plant and animal cells, comprising the nucleic acid sequences, as well as methods for producing and using the proteases, particularly using the proteases in animal feed.

Background of the Invention

In the use of proteases in animal feed (*in vivo*), and/or the use of such proteases for treating vegetable proteins (*in vitro*) it is noted that proteins are essential nutritional factors for animals and humans. Humans and livestock usually get the necessary proteins from vegetable protein sources. Important vegetable protein sources are e.g. oilseed crops, legumes and cereals.

When e.g. soybean meal is included in the feed of mono-gastric animals such as pigs and poultry, a significant proportion of the soybean meal is not digested efficiently (the apparent ileal protein digestibility in piglets, growing pigs and poultry such as broilers, laying hens and roosters is only around 80%).

The gastrointestinal tract of animals consists of a series of segments each representing different pH environments. In mono-gastric animals such as pigs and poultry and many types of fish, the stomach is strongly acidic with a pH potentially as low as 1-2, while the intestine has a more neutral pH of around 6-7. Apart from the stomach and intestine, poultry also have a crop preceding the stomach. The pH in the crop is mostly determined by the feed ingested and hence typically lies in the range of pH 4-6. Protein digestion by a protease may occur along the entire digestive tract, provided that the protease is active and survives the conditions in the digestive tract.

Hence, proteases which are highly acid stable and so can survive in the gastric environment and at the same time are efficiently active at the broad range of physiological pH of the digestive tract in the target animal are especially desirable.

5 Since animal feed is often formulated in pelleted form, in which steam is applied in the pelleting process, it is also desirable that proteases used in animal feed are capable of remaining active after exposure to said steam treatment.

In order to produce a protease for industrial use, it is important that the protease is produced in high yields making the product available in sufficient quantities in order to be able to provide the protease at a favourable price.

10 Description of the Related Art

Proteases of the S1 group isolated from *Saccharopolyspora* are known in the art. Oliynyk *et al* have reported a serine protease from *Saccharopolyspora erythraea* in 'Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338', 2007, *Nat. Biotechnol.* 25:447-453 which has been submitted to EMBL/GenBank (accession number EMBL: AM420293, SEQ ID NO: 1 herein). The amino acid sequence is registered with Uniprot number A4F726 (SEQ ID NO: 2 herein) and the mature amino acid sequence is disclosed in SEQ ID NO: 5.

20 Lucas *et al* have submitted a protease from *Saccharomonospora xinjiangensis* XJ-54 (Uniprot: I0V8H8, SEQ ID NO: 8) to EMBL/Genbank having 76.5% homology to sequence identity to SEQ ID NO: 5. Lucas *et al* have submitted an endopeptidase from *Saccharomonospora paurometabolica* YIM 90007 (Uniprot: G4J6Q2, SEQ ID NO: 9) having 74.7% sequence identity to SEQ ID NO: 5 to EMBL/Genbank.

25 Lucas *et al* have submitted the Chymotrypsin-like protease from *Kribbella flavida* (Uniprot: D2PRB9, SEQ ID NO: 10) having 74.7% homology to sequence identity to SEQ ID NO: 5 to EMBL/Genbank. Pati *et al.* have disclosed a serine protease from *Saccharomonospora viridis* in "Complete genome sequence of *Saccharomonospora viridis* type strain (P101)", 2009, *Stand. Genomic Sci.* 1:141-149 (Uniprot: C7MV18, SEQ ID NO: 11) with 73.8% sequence homology to SEQ ID NO: 5. Strobel *et al* have published 2 proteases from *Saccharothrix espanaensis* DSM 44229 having 74.3% and 73.6% sequence identity to to SEQ ID NO: 5 (Uniprot: K0JWC2, SEQ ID NO: 12 and K0JQQ4, SEQ ID NO: 13 respectively) in 'Complete genome sequence of *Saccharothrix espanaensis* DSM 44229T and comparison to the other completely sequenced Pseudonocardiaceae', 2012, *BMC Genomics*, 13:465-465.

The next closest *Saccharopolyspora* protease identified was submitted to EMBL/Genbank (Uniprot A4FNQ0) with 69.6 % sequence identity to SEQ ID NO: 5. Other enzymes have less than 74% sequence identity to SEQ ID NO: 5.

5 WO 05/052146 and WO 05/052161 describe a number of serine proteases used for animal feed having an identity to the protease of SEQ ID NO: 5 of 67-73 %.

10 WO 95/28850 discloses the combination of a phytase and one or more microbial proteolytic enzymes to improve the solubility of vegetable proteins. WO 01/58275 discloses the use of acid stable proteases of the subtilisin family in animal feed. WO 01/58276 discloses the use of acid-stable proteases derived from *Nocardioopsis sp.* NRRL 18262 (a 10R protease), as well as a protease derived from *Nocardioopsis alba* DSM 14010 in animal feed. WO 04/072221, WO 04/111220, WO 04/111223, WO 05/035747, and WO 05/123911 disclose proteases related to the 10R protease and their use in animal feed. WO 04/072279 discloses the use of other proteases in animal feed. WO 04/034776 discloses the use of a subtilisin/keratinase, PWD-1 from *B. Licheniformis*, in the feed of poultry. WO 04/077960 discloses a method for increasing the 15 digestibility of forage or grain in ruminants by applying a bacterial or fungal protease.

Commercial products comprising a protease and marketed for use in animal feed include RONOZYME® ProAct (DSM NP/Novozymes), Aextra® (Danisco), Avizyme® (Danisco), Porzyme® (Danisco), Allzyme™ (Alltech), Versazyme® (BioResources, Int.), Poultrygrow™ (Jefo) and Cibenza® DP100 (Novus).

20

SUMMARY OF THE INVENTION

The present invention relates to isolated polypeptides having protease activity for use in animal feed selected from the group consisting of:

- (a) a polypeptide having at least 80% sequence identity to the polypeptide of SEQ ID NO: 5;
- 25 (b) a polypeptide encoded by a polynucleotide that hybridizes under medium-high stringency conditions, high stringency conditions or very high stringency conditions with:
- (i) the mature polypeptide coding sequence of SEQ ID NO: 1;
 - (ii) the mature polypeptide coding sequence of SEQ ID NO: 3; and/or
 - (iii) the full-length complementary strand of (i) or (ii);
- 30 (c) a polypeptide encoded by a polynucleotide having at least 80% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or SEQ ID NO: 3;

(d) a variant of the polypeptide of SEQ ID NO: 5 having at least 80% sequence identity to the polypeptide of SEQ ID NO: 5 comprising a substitution, deletion, and/or insertion at one or more (several) positions; and

(e) a fragment of a polypeptide of (a), (b), (c) or (d) that has protease activity.

5

The present invention also relates to variant polypeptides having protease activity and having at least 80% sequence identity to the polypeptide of SEQ ID NO: 5, selected from the group consisting of:

(a) a variant of the polypeptide of SEQ ID NO: 5 comprising at least one substitution, deletion, and/or insertion at one or more (several) positions; and

10

(b) a fragment of a polypeptide of (a) that has protease activity.

15

The present invention relates to isolated polynucleotides encoding the polypeptides of the present invention, nucleic acid constructs, recombinant expression vectors, and recombinant host cells comprising the polynucleotides, and to methods of producing the polypeptides.

The present invention also relates to compositions of the isolated polypeptide of the present invention, methods for preparing a composition for use in animal feed, improving the nutritional value of an animal feed, and methods of treating proteins to be used in animal feed compositions.

Overview of Sequence Listing

20

SEQ ID NO: 1 is the DNA sequence of S1 protease 2 as isolated from *Saccharopolyspora erythraea* (EMBL: AM420293).

SEQ ID NO: 2 is the amino acid sequence as deduced from SEQ ID NO: 1.

SEQ ID NO: 3 is the DNA sequence of the synthetically optimised gene fused with the Savinase signal peptide.

25

SEQ ID NO: 4 is the amino acid sequence as deduced from SEQ ID NO: 3.

SEQ ID NO: 5 is the amino acid sequence of the mature S1 protease 2 from *Saccharopolyspora erythraea*.

SEQ ID NO: 6 is the DNA sequence of protease 10R (WO 05/035747, SEQ ID NO: 1).

SEQ ID NO: 7 is the amino acid sequence of protease 10R (WO 05/035747, SEQ ID NO:

30

2).

SEQ ID NO: 8 is the amino acid sequence of the protease from *Saccharomonospora xinjiangensis* XJ-54 (Uniprot: I0V8H8).

SEQ ID NO: 9 is the amino acid sequence of an endopeptidase from *Saccharomonospora paurometabolica* YIM 90007 (Uniprot: G4J6Q2).

SEQ ID NO: 10 is the amino acid sequence of a Chymotrypsin-like protease from *Kribbella flavida* DSM 17836 (Uniprot: D2PRB9).

5 SEQ ID NO: 11 is the amino acid sequence of a serine protease from *Saccharomonospora viridis* (Uniprot: C7MV18).

SEQ ID NO: 12 is the amino acid sequence of a Streptogrisin C like protein *Saccharothrix espanaensis* DSM 44229T (Uniprot: K0JWC2).

10 SEQ ID NO: 13 is the amino acid sequence of a Streptogrisin C like protein *Saccharothrix espanaensis* DSM 44229T (Uniprot: K0JQQ4).

SEQ ID NO: 14 is the *Bacillus lentus* secretion signal.

Identity Matrix of sequences:

	SEQ ID NO: 2	SEQ ID NO: 4	SEQ ID NO: 5	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 9	SEQ ID NO: 10	SEQ ID NO: 11	SEQ ID NO: 12	SEQ ID NO: 13
SEQ2	100	100	100	48.8	66.6	62.6	58.7	64.5	58.4	57.2
SEQ4	100	100	100	48.8	66.6	62.6	58.7	64.5	58.4	57.2
SEQ5	100	100	100	58.9	76.5	74.7	74.7	73.8	74.3	73.6
SEQ7	48.8	48.8	58.9	100	49.7	48.5	47.8	50.0	48.5	43.6
SEQ8	66.6	66.6	76.5	49.7	100	73.1	54.6	73.8	55.2	54.5
SEQ9	62.6	62.6	74.7	48.5	73.1	100	53.4	70.1	55.2	52.0
SEQ10	58.7	58.7	74.7	47.8	54.6	53.4	100	55.5	55.2	66.0
SEQ11	64.5	64.5	73.8	50.0	73.8	70.1	55.5	100	55.8	52.9
SEQ12	58.4	58.4	74.3	48.5	55.2	55.2	55.2	55.8	100	57.2
SEQ13	57.2	57.2	73.6	43.6	54.5	52.0	66.0	52.9	57.2	100

Brief Description of the Figures

15 Figure 1 shows the pH-activity profile of S1 protease 2 as isolated from *Saccharopolyspora erythraea* compared to protease 10R on the Suc-AAPF-pNA substrate at 37°C.

Figure 2 shows the pH-stability profile of S1 protease 2 as isolated from *Saccharopolyspora erythraea* compared to protease 10R (residual activity after 2 hours at 37°C).

Figure 3 shows the temperature activity profile of S1 protease 2 as isolated from *Saccharopolyspora erythraea* at pH 7.0 compared to protease 10R on Protazyme AK at pH 6.5.

5 Figure 4 shows the P1-specificity of S1 protease 2 as isolated from *Saccharopolyspora erythraea* compared to protease 10R on 10 Suc-AAPX-pNA substrates at pH 9.0, 25°C.

Figure 5 shows the activity ($OD_{340} \times$ dilution factor) on soybean-maize meal of the S1 protease 2 as isolated from *Saccharopolyspora erythraea* compared to protease 10R.

DEFINITIONS

10 **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.

15 **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 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.

20 **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.

25 **Control sequences:** The term "control sequences" means nucleic acid sequences necessary for expression of a polynucleotide encoding a mature polypeptide 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 polypeptide or native or foreign to each other. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, 30 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 polypeptide.

Expression: The term "expression" includes any step involved in the production of the polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, 5 post-translational modification, and secretion.

Expression vector: The term "expression vector" means a linear or circular DNA molecule that comprises a polynucleotide encoding a polypeptide and is operably linked to additional nucleotides that provide for its expression.

Fragment: The term "fragment" means a polypeptide having one or more (several) amino 10 acids deleted from the amino and/or carboxyl terminus of a mature polypeptide; wherein the fragment has protease activity. In one aspect, a fragment contains at least 163 amino acid residues (e.g., amino acids 11 to 173 of SEQ ID NO: 2), or at least 173 amino acid residues (e.g., amino acids 6 to 178 of SEQ ID NO: 2); or correspondingly for SEQ ID NO: 4 a fragment contains at least 163 amino acid residues (e.g., amino acids 11 to 173 of SEQ ID NO: 4), or at least 173 amino acid 15 residues (e.g., amino acids 6 to 178 of SEQ ID NO: 4); or correspondingly for SEQ ID NO: 5 a fragment contains at least 163 amino acid residues (e.g., amino acids 11 to 173 of SEQ ID NO: 5), or at least 173 amino acid residues (e.g., amino acids 6 to 178 of SEQ ID NO: 5).

Host cell: The term "host cell" means any cell type that is susceptible to transformation, transfection, transduction, and the like with a nucleic acid construct or expression vector comprising 20 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.

Isolated polynucleotide: The term "isolated polynucleotide" means a polynucleotide that is modified by the hand of man relative to that polynucleotide as found in nature. In one aspect, the isolated polynucleotide is at least 1% pure, e.g., at least 5% pure, more at least 10% pure, at least 25 20% pure, at least 40% pure, at least 60% pure, at least 80% pure, at least 90% pure, and at least 95% pure, as determined by agarose electrophoresis. The polynucleotides may be of genomic, cDNA, RNA, semisynthetic, synthetic origin, or any combinations thereof.

Isolated polypeptide: The term "isolated polypeptide" means a polypeptide that is modified by the hand of man relative to that polypeptide as found in nature. In one aspect, the polypeptide is 30 at least 1% pure, e.g., at least 5% pure, at least 10% pure, at least 20% pure, at least 40% pure, at least 60% pure, at least 80% pure, and at least 90% pure, as determined by SDS-PAGE.

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 35 amino acids 1 to 183 of SEQ ID NO: 2 based on the SignalP (Nielsen *et al.*, 1997, *Protein*

5 *Engineering* 10: 1-6) prediction program that also predicts -194 to -154 of SEQ ID NO: 2 are a signal peptide. In another aspect, the mature polypeptide is amino acids 1 to 183 of SEQ ID NO: 4 based on sequencing using Edman degradation and intact molecular weight analysis. Amino acids -180 to -154 of SEQ ID NO: 4 is the Savinase signal peptide. It is known in the art that a host cell may produce a mixture of two or more different mature polypeptides (*i.e.*, with a different C-terminal and/or N-terminal amino acid) expressed by the same polynucleotide. It is also known in the art that different host cells process polypeptides differently, and thus, one host cell expressing a polynucleotide may produce a different mature polypeptide (*e.g.*, having a different C-terminal and/or N-terminal amino acid) as compared to another host cell expressing the same
10 polynucleotide.

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 683 to 1231 in the numbering of SEQ ID NO: 1 and nucleotides 541 to 1089 in the numbering of SEQ ID NO: 3 based on the determination
15 of the mature polypeptide by Edman degradation and intact molecular weight analysis. Furthermore nucleotides 101 to 223 in the numbering of SEQ ID NO: 1 are predicted to encode a signal peptide based on the prediction program SignalP (Nielsen *et al.*, 1997, *Protein Engineering* 10: 1-6) and nucleotides 1 to 81 in the numbering of SEQ ID NO: 3 encode the Savinase signal peptide.

Nucleic acid construct: The term "nucleic acid construct" means a nucleic acid molecule,
20 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. The term nucleic acid construct is synonymous with the term "expression cassette" when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present invention.

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 the expression of the coding sequence.
25

Protease activity: The term "protease activity" means proteolytic activity (EC 3.4). Proteases of the invention are endopeptidases (EC 3.4.21). There are several protease activity
30 types, although the three main activity types are: trypsin-like where there is cleavage of amide substrates following Arg or Lys at P1, chymotrypsin-like where cleavage occurs following one of the hydrophobic amino acids at P1, and elastase-like with cleavage following an Ala at P1.

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-
35 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).

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.

The polypeptides 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%, and at least 100% of the protease activity of the polypeptide of SEQ ID NO: 5.

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

For purposes of the present invention, the degree of 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 3.0.0 or later. Version 6.1.0 was used. The optional 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:

$$\text{(Identical Residues x 100)} / (\text{Length of Alignment} - \text{Total Number of Gaps in Alignment})$$

For purposes of the present invention, the degree of 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 3.0.0 or later. Version 6.1.0 was used. The optional 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:

$$\frac{(\text{Identical Deoxyribonucleotides} \times 100)}{(\text{Length of Alignment} - \text{Total Number of Gaps in Alignment})}$$

5 **Stringency conditions:** The different stringency conditions are defined as follows.

10 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.

15 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.

20 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.

25 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.

30 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.

35 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.

Subsequence: The term “subsequence” means a polynucleotide having one or more (several) nucleotides deleted 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 489 nucleotides (*e.g.*, nucleotides 713 to 1201 of SEQ ID NO: 1 or nucleotides 571 to 1059 of SEQ ID NO: 3). In another aspect, a subsequence contains at least 519 nucleotides (*e.g.*, nucleotides 698 to 1216 of SEQ ID NO: 1 or nucleotides 556 to 1074 of SEQ ID NO: 3).

Substantially pure polynucleotide: The term “substantially pure polynucleotide” means a polynucleotide preparation free of other extraneous or unwanted nucleotides and in a form suitable for use within genetically engineered polypeptide production systems. Thus, a substantially pure polynucleotide contains at most 10%, at most 8%, at most 6%, at most 5%, at most 4%, at most 3%, at most 2%, at most 1%, and at most 0.5% by weight of other polynucleotide material with which it is natively or recombinantly associated. A substantially pure polynucleotide may, however, include naturally occurring 5' and 3' untranslated regions, such as promoters and terminators. Preferably, the polynucleotide is at least 90% pure, *e.g.*, at least 92% pure, at least 94% pure, at least 95% pure, at least 96% pure, at least 97% pure, at least 98% pure, at least 99% pure, and at least 99.5% pure by weight. The polynucleotides of the present invention are preferably in a substantially pure form.

Substantially pure polypeptide: The term “substantially pure polypeptide” means a preparation that contains at most 10%, at most 8%, at most 6%, at most 5%, at most 4%, at most 3%, at most 2%, at most 1%, and at most 0.5% by weight of other polypeptide material with which it is natively or recombinantly associated. Preferably, the polypeptide is at least 92% pure, *e.g.*, at least 94% pure, at least 95% pure, at least 96% pure, at least 97% pure, at least 98% pure, at least 99%, at least 99.5% pure, and 100% pure by weight of the total polypeptide material present in the preparation. The polypeptides of the present invention are preferably in a substantially pure form. This can be accomplished, for example, by preparing the polypeptide by well known recombinant methods or by classical purification methods.

Variant: The term “variant” means a polypeptide having protease activity and having at least 80%, *e.g.*, 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%, at least 99% sequence identity to SEQ ID NO: 5, comprising an alteration, *i.e.*, a substitution, insertion, and/or deletion of one or more (several) amino acid residues at one or more (several) positions. A substitution means a replacement of an amino acid occupying a position with a different amino acid; a deletion means removal of *e.g.* 1-5 amino acid residues occupying 1-5 positions; and an insertion means adding *e.g.* 1-5 amino acids adjacent to an amino acid occupying

a position. 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 polypeptide of SEQ ID NO: 5. A variant may also be a naturally occurring protease having at least 80%, *e.g.*, 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 at least 99% sequence identity to SEQ ID NO: 5.

DETAILED DESCRIPTION OF THE INVENTION

Polypeptides Having Protease Activity

Polypeptides having protease activity, or proteases, are sometimes also designated peptidases, proteinases, peptide hydrolases, or proteolytic enzymes. Proteases may be of the exo-type that hydrolyse peptides starting at either end thereof, or of the endo-type that act internally in polypeptide chains (endopeptidases). Endopeptidases show activity on N- and C-terminally blocked peptide substrates that are relevant for the specificity of the protease in question.

The term "protease" is defined herein as an enzyme that hydrolyses peptide bonds. This definition of protease also applies to the protease-part of the terms "parent protease" and "protease variant," as used herein. The term "protease" includes any enzyme belonging to the EC 3.4 enzyme group (including each of the eighteen subclasses thereof). The EC number refers to Enzyme Nomenclature 1992 from NC-IUBMB, Academic Press, San Diego, California, including supplements 1-5 published in 1994, *Eur. J. Biochem.* 223: 1-5; 1995, *Eur. J. Biochem.* 232: 1-6; 1996, *Eur. J. Biochem.* 237: 1-5; 1997, *Eur. J. Biochem.* 250: 1-6; and 1999, *Eur. J. Biochem.* 264: 610-650 respectively. The nomenclature is regularly supplemented and updated; see *e.g.* the World Wide Web (WWW) at <http://www.chem.qmw.ac.uk/iubmb/enzyme/index.html>.

The present invention provides for the use of polypeptides having protease activity in animal feed compositions. It also provides polynucleotides encoding the polypeptides. The proteases of the invention are serine proteases of the peptidase family S1. The proteases of the invention exhibit surprising pH properties, which makes them interesting candidates for use in animal feed. The proteases of the invention are thus most active on Suc-Ala-Ala-Pro-Phe-pNA but have reasonable activity on Suc-Ala-Ala-Pro-Met-pNA within a broad pH range of 5-11 and exhibit especially high activity in the pH range of 6-11. They are active on a feed relevant soybean meal-maize meal substrate within a broad physiological pH range of pH 3-7 and retain more than 90% activity after being subjected for 2 hours to a pH as low as 4.

The proteases of the invention and for use according to the invention are selected from the group consisting of:

(a) proteases belonging to the EC 3.4.21. enzyme group; and/or

(b) Serine proteases of the peptidase family S1;

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.

For determining whether a given protease is a serine protease, and a family S1 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.

Proteases of the invention are endopeptidases (EC 3.4.21). There are several protease activity types: The three main activity types are: trypsin-like where there is cleavage of amide substrates following Arg or Lys at P1, chymotrypsin-like where cleavage occurs following one of the hydrophobic amino acids at P1, and elastase-like with cleavage following an Ala at P1.

The peptidases of family S1 contain the catalytic triad in the order His, Asp, Ser. Mutation of any of the amino acids of the catalytic triad will result in change or loss of enzyme activity. The amino acids of the catalytic triad of the S1 protease 2 as isolated from *Saccharopolyspora erythraea* (SEQ ID NO: 5) are probably positions His-32, Asp-56 and Ser-137.

The polypeptides 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%, at least 96%, at least 97%, at least 98%, at least 99%, and at least 100% of the protease activity of the mature polypeptide of SEQ ID NO: 5.

The present invention provides polypeptides having protease activity and polynucleotides encoding the polypeptides. The proteases of the invention are serine proteases of the peptidase family S1. The proteases of the invention exhibit pH properties, especially pH stability properties which make them of substantial interest as candidates for use in animal feed and other applications.

The present invention relates to the use of isolated polypeptides having protease activity in animal feed selected from the group consisting of:

(a) a polypeptide having 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% sequence identity to the polypeptide of SEQ ID NO: 5;

(b) a polypeptide encoded by a polynucleotide that hybridizes under medium-high stringency conditions, high stringency conditions, or very high stringency conditions with:

(i) the mature polypeptide coding sequence of SEQ ID NO: 1;

- (ii) the mature polypeptide coding sequence of SEQ ID NO: 3;
- (iii) the full-length complementary strand of (i) or (ii);
- (c) a polypeptide encoded by a polynucleotide having 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% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or SEQ ID NO: 3;
- (d) a variant of the polypeptide of SEQ ID NO: 5 having at least 80% sequence identity to the polypeptide of SEQ ID NO: 5 comprising a substitution, deletion, and/or insertion at one or more (several) positions; and
- (e) a fragment of a polypeptide of (a), (b), (c) or (d) that has protease activity.

The present invention relates to the use in animal feed isolated polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 80%, *e.g.*, 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%, at least 99%, or 100%, which have protease activity. In one aspect, the polypeptides differ by no more than thirtysix amino acids, *e.g.*, by thirty amino acids, by twentyfive amino acids, by twenty amino acids, by fifteen amino acids, by ten amino acids, by nine amino acids, by eight amino acids, by seven amino acids, by six amino acids, by five amino acids, by four amino acids, by three amino acids, by two amino acids, and by one amino acid from the mature polypeptide of SEQ ID NO: 2.

The present invention also relates to the use in animal feed of isolated polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 4 of at least 80%, *e.g.*, 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%, at least 99%, or 100%, which have protease activity. In one aspect, the polypeptides differ by no than thirtysix amino acids, *e.g.*, by thirty amino acids, by twentyfive amino acids, by twenty amino acids, by fifteen amino acids, by ten amino acids, by nine amino acids, by eight amino acids, by seven amino acids, by six amino acids, by five amino acids, by four amino acids, by three amino acids, by two amino acids, and by one amino acid from the mature polypeptide of SEQ ID NO: 4.

The present invention further relates to the use in animal feed of isolated polypeptides having a sequence identity to the polypeptide of SEQ ID NO: 5 of at least 80%, *e.g.*, 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%, at least 99%, or 100%, which have protease activity. In one aspect, the polypeptides differ by no more than thirtysix amino acids, *e.g.*, by thirty amino acids, by twentyfive amino acids, by twenty amino acids, by

fifteen amino acids, by ten amino acids, by nine amino acids, by eight amino acids, by seven amino acids, by six amino acids, by five amino acids, by four amino acids, by three amino acids, by two amino acids, and by one amino acid from the mature polypeptide of SEQ ID NO: 5.

5 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 85% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 86% sequence identity to the polypeptide of SEQ ID NO: 5.

10 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 87% sequence identity to the polypeptide of SEQ ID NO: 5.

15 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 88% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 89% sequence identity to the polypeptide of SEQ ID NO: 5.

20 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 90% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 91% sequence identity to the polypeptide of SEQ ID NO: 5.

25 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 92% sequence identity to the polypeptide of SEQ ID NO: 5.

30 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 93% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 94% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 95% sequence identity to the polypeptide of SEQ ID NO: 5.

5 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 96% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 97% sequence identity to the polypeptide of SEQ ID NO: 5.

10 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 98% sequence identity to the polypeptide of SEQ ID NO: 5.

15 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having at least 99% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide for use in animal feed having 100% sequence identity to the polypeptide of SEQ ID NO: 5.

20 The present invention also relates to isolated polypeptides having protease activity and having at least 80% sequence identity to the polypeptide of SEQ ID NO: 5, selected from the group consisting of:

(a) a variant of of the polypeptide of SEQ ID NO: 5 comprising a substitution, deletion, and/or insertion at one or more (several) positions; and

(b) a fragment of a polypeptide of (a) that has protease activity.

25 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 85% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 86% sequence identity to the polypeptide of SEQ ID NO: 5.

30 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 87% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 88% sequence identity to the polypeptide of SEQ ID NO: 5.

35 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 89% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 90% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 91% sequence identity to the polypeptide of SEQ ID NO: 5.

5 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 92% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 93% sequence identity to the polypeptide of SEQ ID NO: 5.

10 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 94% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 95% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 96% sequence identity to the polypeptide of SEQ ID NO: 5.

15 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 97% sequence identity to the polypeptide of SEQ ID NO: 5.

An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 98% sequence identity to the polypeptide of SEQ ID NO: 5.

20 An embodiment of the invention is a polypeptide or a polypeptide encoded by a polynucleotide having at least 99% sequence identity to the polypeptide of SEQ ID NO: 5.

A polypeptide of the present invention preferably comprises or consists of the amino acid sequence of SEQ ID NO: 2, SEQ ID NO: 4 or SEQ ID NO: 5, an allelic variant thereof; or is a fragment missing e.g. 30, 25, 20, 15, 10 or 5 amino acids from the *N*- and/or *C*-terminal and having
25 protease activity. In another aspect, the polypeptide comprises or consists of the polypeptide of SEQ ID NO: 2, SEQ ID NO: 4 and/or SEQ ID NO: 5. In another preferred aspect, the polypeptide comprises or consists of amino acids 1 to 193 of SEQ ID NO: 2, amino acids 1 to 193 of SEQ ID NO: 4, and/or amino acids 1 to 193 of SEQ ID NO: 5.

30 The present invention also relates to isolated polypeptides having protease activity that are encoded by polynucleotides that hybridize under medium-high stringency conditions, high stringency conditions, or very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, (ii) the mature polypeptide coding sequence of SEQ ID NO: 3, or (iii) the full-length complementary strand of (i) or (ii) (J. Sambrook, E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, New York).

The polynucleotide of SEQ ID NO: 1, SEQ ID NO: 3 or a subsequence thereof, as well as the amino acid sequence of SEQ ID NO: 2, SEQ ID NO: 4 or SEQ ID NO: 5, or a fragment thereof, may be used to design nucleic acid probes to identify and clone DNA encoding polypeptides having protease activity 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 or cDNA of the genus or species 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 14, *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 ³²P, ³H, ³⁵S, biotin, or avidin). Such probes are encompassed by the present invention.

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 polypeptide having protease activity. 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 is homologous with SEQ ID NO: 1, SEQ ID NO: 3 or a subsequence thereof, the carrier material is preferably used in a Southern blot.

For purposes of the present invention, hybridization indicates that the polynucleotide hybridizes to a labeled nucleic acid probe corresponding to the mature polypeptide coding sequence of SEQ ID NO: 1, the mature polypeptide coding sequence of SEQ ID NO: 3; its full-length complementary strand; or a subsequence thereof; under very low 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.

In one aspect, the nucleic acid probe is the mature polypeptide coding sequence of SEQ ID NO: 1 or SEQ ID NO: 3. In another aspect, the nucleic acid probe is a fragment thereof. In another aspect, the nucleic acid probe is a polynucleotide that encodes the polypeptide of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5 or a fragment thereof. In another preferred aspect, the nucleic acid probe is SEQ ID NO: 1 or SEQ ID NO: 3.

For long probes of at least 100 nucleotides in length, high to very high stringency conditions are defined as prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and either 25% formamide for very low

and low stringencies, 35% formamide for medium and medium-high stringencies, or 50% formamide for high and very high stringencies, following standard Southern blotting procedures for 12 to 24 hours optimally. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 65°C (high stringency), and at 70°C (very high stringency).

5 For short probes of about 15 nucleotides to about 70 nucleotides in length, stringency conditions are defined as prehybridization and hybridization at about 5°C to about 10°C below the calculated T_m using the calculation according to Bolton and McCarthy (1962, *Proc. Natl. Acad. Sci. USA* 48:1390) in 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1X Denhardt's solution, 1 mM sodium pyrophosphate, 1 mM sodium monobasic phosphate, 0.1 mM ATP, and 0.2
10 mg of yeast RNA per ml following standard Southern blotting procedures for 12 to 24 hours optimally. The carrier material is finally washed once in 6X SSC plus 0.1% SDS for 15 minutes and twice each for 15 minutes using 6X SSC at 5°C to 10°C below the calculated T_m .

The present invention also relates to the use in animal feed of isolated polypeptides having protease activity encoded by polynucleotides having a sequence identity to the mature polypeptide
15 coding sequence of SEQ ID NO: 1 of at least least 80%, e.g., 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%, at least 99%, or 100%.

The present invention also relates to variant polypeptides having protease activity and having at least 80%, e.g., at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at
20 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% sequence identity to the polypeptide of SEQ ID NO: 5 comprising at least one substitution, deletion, and/or insertion of at least one or more (several) amino acids of SEQ ID NO: 5 or a homologous sequence thereof.

The variant polypeptide of the invention may in one embodiment have at least 86%
25 sequence identity to SEQ ID NO: 5.

The variant polypeptide of the invention may in one embodiment have at least 87% sequence identity to SEQ ID NO: 5.

The variant polypeptide of the invention may in one embodiment have at least 88% sequence identity to SEQ ID NO: 5.

30 The variant polypeptide of the invention may in one embodiment have at least 89% sequence identity to SEQ ID NO: 5.

The variant polypeptide of the invention may in one embodiment have at least 90% sequence identity to SEQ ID NO: 5.

35 The variant polypeptide of the invention may in one embodiment have at least 91% sequence identity to SEQ ID NO: 5.

The variant polypeptide of the invention may in one embodiment have at least 92% sequence identity to SEQ ID NO: 5.

The variant polypeptide of the invention may in one embodiment have at least 93% sequence identity to SEQ ID NO: 5.

5 The variant polypeptide of the invention may in one embodiment have at least 94% sequence identity to SEQ ID NO: 5.

The variant polypeptide of the invention may in one embodiment have at least 95% sequence identity to SEQ ID NO: 5.

10 The variant polypeptide of the invention may in one embodiment have at least 96% sequence identity to SEQ ID NO: 5.

The variant polypeptide of the invention may in one embodiment have at least 97% sequence identity to SEQ ID NO: 5.

The variant polypeptide of the invention may in one embodiment have at least 98% sequence identity to SEQ ID NO: 5.

15 The variant polypeptide of the invention may in one embodiment have at least 99% sequence identity to SEQ ID NO: 5.

In a further embodiment, the total number of positions of the variant polypeptide of the invention (SEQ ID NO: 5) having amino acid substitutions, deletions and/or insertions is not more than 36, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 20 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 or 36. 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 one to about 30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to about 20-25 residues; or a small extension that facilitates purification by 25 changing net charge or another function, such as a poly-histidine tract, an antigenic epitope or a binding domain.

The present invention also relates to the use in animal feed of isolated polypeptides having protease activity encoded by polynucleotides having a sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3 of at least least 80%, e.g., at least 85%, at least 86%, at least 30 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%, at least 99%, or 100%

In another embodiment, the present invention also relates to variants for use in animal feed comprising a substitution, deletion, and/or insertion of one or more (or several) amino acids of SEQ ID NO: 2 or a homologous sequence thereof. The total number of positions having amino acid 35 substitutions, deletions and/or insertions in SEQ ID NO: 2 is not more than 36, e.g., 1, 2, 3, 4, 5, 6,

7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 or 36. 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 one to about 30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to about 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.

In another embodiment, the present invention also relates to variants for use in animal feed comprising a substitution, deletion, and/or insertion of one or more (or several) amino acids of SEQ ID NO: 4 or a homologous sequence thereof. The total number of positions having amino acid substitutions, deletions and/or insertions in SEQ ID NO: 4 is not more than 36, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 or 36. 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 one to about 30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to about 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.

In another embodiment, the present invention also relates to variants for use in animal feed comprising a substitution, deletion, and/or insertion of one or more (or several) amino acids of SEQ ID NO: 5 or a homologous sequence thereof. The total number of positions having amino acid substitutions, deletions and/or insertions in SEQ ID NO: 5 is not more than 36, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 or 36. 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 one to about 30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to about 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.

Examples of conservative substitutions are within the group 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. The most commonly occurring exchanges that are expected not to alter the specific activity substantially 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.

5 Alternatively, the amino acid changes are of such a nature that the physico-chemical properties of the polypeptides are altered. For example, amino acid changes may improve the thermal stability of the polypeptide, alter the substrate specificity, change the pH optimum, and the like.

10 Essential amino acids in a parent 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
15 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. The identities of essential amino acids can also be
20 inferred from analysis of identities with polypeptides that are related to the parent polypeptide.

 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;
25 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).

30 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.

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

The polypeptide 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 fused 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 fused polypeptide is under control of the same promoter(s) and terminator. Fusion proteins may also be constructed using intein technology in which fusions are created post-translationally (Cooper *et al.*, 1993, *EMBO J.* 12: 2575-2583; Dawson *et al.*, 1994, *Science* 266: 776-779).

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.

20 Embodiments

In certain embodiments of the invention, the protease of the invention exhibits beneficial thermal properties such as thermostability, steam stability, etc and/or pH properties, such as acid stability, pH optimum, etc.

An embodiment of the invention is isolated polypeptides for use in animal feed having improved protease activity between pH 6 and 9, at 37°C compared to protease 10R.

A further embodiment of the invention is isolated polypeptides for use in animal feed having improved protease activity between 15°C and 60°C, such as between 25°C and 50°C, or at 15°C, at 25°C, at 37°C, at 50°C or at 60°C compared to protease 10R.

An additional embodiment of the invention is isolated polypeptides for use in animal feed having improved protease activity on soybean-maze meal between pH 3.0 and 5.0, such as at pH 3.0, 4.0 or 5.0 at 40°C compared to protease 10R.

Acidity/alkalinity properties

In certain embodiments of the invention the protease of the invention exhibits beneficial properties in respect of pH, such as acid stability, pH optimum, etc. Stability of the protease at a low pH is beneficial since the protease can have activity in the intestine after passing through the stomach. In one embodiment of the invention the protease retains >90% activity after 2 hours at pH 4 as determined using the method described in Example 3.

Temperature-activity

The temperature-activity profile of the protease may be determined as described in Example 3. Activity at low temperatures (30-40°C) can be advantageous for the digestion of proteins in an animal.

In one embodiment, the invention comprises of a protease having a temperature activity profile at pH 7.0 with relative activity of 0.15 or higher at 25°C, or relative activity of 0.40 or higher at 37°C, or relative activity of 0.75 or higher at 50°C when compared to the activity of the protease at 60°C (cf. Example 3).

pH-activity

The pH-activity profile of the protease may be determined as described in Example 3. Activity at pH 6-8 can be advantageous for the digestion of proteins in the intestine of an animal.

In one embodiment, the invention comprises of a protease for use in animal feed having a pH-activity profile at 37°C with relative activity of 0.50 or higher at pH 7, or relative activity of 0.80 or higher at pH 8 when compared to the activity of the protease at pH 9 (cf. Example 3).

Thermostability

Thermostability may be determined as described in Example 5, i.e. using DSC measurements to determine the denaturation temperature, T_d , of the purified protease protein. The T_d is indicative of the thermostability of the protein: The higher the T_d , the higher the thermostability. Accordingly, in a preferred embodiment, the protease of the invention has a T_d which is higher than the T_d of a reference protease, wherein T_d is determined on purified protease samples (preferably with a purity of at least 90% or 95%, as determined by SDS-PAGE).

In preferred embodiments, the thermal properties such as heat-stability, temperature stability, thermostability, steam stability, and/or pelleting stability as provided by the residual activity, denaturation temperature T_d , or other parameter of the protease of the invention is higher than the corresponding value, such as the residual activity or T_d , of the protease of SEQ ID NO: 5, more preferably at least 101% thereof, or at least 102%, 103%, 104%, 105%, 106%, 107%, 108%, 109%, or at least 110% thereof. Even more preferably, the value of the parameter, such as residual

activity or T_d , of the protease of the invention is at least 120%, 130%, 140%, 150%, 160%, 170%, 180%, or at least 190% of the value for the protease of SEQ ID NO: 5.

In still further particular embodiments, the thermostable protease of the invention has a melting temperature, T_m (or a denaturation temperature, T_d), as determined using Differential Scanning Calorimetry (DSC) as described in example 10 (i.e. in 20 mM sodium acetate, pH 4.0), of at least 50°C. In still further particular embodiments, the T_m is at least 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or at least 100°C.

Steam stability

10 Steam stability may be determined as described in Example 6 by determining the residual activity of protease molecules after steam treatment at 85°C or 90°C for a short time.

Pelleting stability

15 Pelleting stability may be determined as described in Example 7 by using enzyme granulate pre-mixed with feed. From the mixer the feed is conditioned with steam to 95°C. After conditioning the feed is pressed to pellets and the residual activity determined.

Sources of Polypeptides Having Protease Activity

20 A polypeptide having protease activity and to be used according to the present invention may be obtained from microorganisms of any genus. For purposes of the present invention, the term "obtained from" as used herein in connection with a given source shall mean that the polypeptide encoded by a polynucleotide is produced by the source or by a strain in which the polynucleotide from the source has been inserted. In one aspect, the polypeptide obtained from a given source is secreted extracellularly.

25 The polypeptide may be a bacterial polypeptide. For example, the polypeptide may be a polypeptide having protease activity from a gram-positive bacterium within a phylum such as Actinobacteria or from a gram-negative bacterium within a phylum such as Proteobacteria.

In one aspect, the polypeptide is a protease from a bacterium of the class *Actinobacteria*, such as from the order *Actinomycetales*, or from the suborder *Pseudonocardineae*, or from the family *Pseudonocardiaceae*, or from the genus *Saccharopolyspora*, or from the species *Saccharopolyspora erythraea*.

30 It will be understood that for the aforementioned species, the invention encompasses both the perfect and imperfect states, and other taxonomic equivalents, e.g., anamorphs, regardless of

the species name by which they are known. Those skilled in the art will readily recognize the identity of appropriate equivalents.

5 Strains of these taxa are readily accessible to the public in a number of culture collections, such as the American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau Voor Schimmelcultures (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

10 The polypeptide may be identified and obtained from other sources including microorganisms isolated from nature (e.g., soil, composts, water, etc.) using the above-mentioned probes. Techniques for isolating microorganisms from natural habitats are well known in the art. The polynucleotide encoding the polypeptide may then be obtained by similarly screening a genomic or cDNA library of another microorganism or mixed DNA sample. Once a polynucleotide encoding a polypeptide has been detected with the probe(s), the polynucleotide can be isolated or cloned by utilizing techniques that are well known to those of ordinary skill in the art (see, e.g., Sambrook *et al.*, 1989, *supra*).

15 **Polynucleotides**

The present invention also relates to isolated polynucleotides encoding a polypeptide of the present invention.

20 The techniques used to isolate or clone a polynucleotide encoding a polypeptide are known in the art and include isolation from genomic DNA, preparation from cDNA, or a combination thereof. The cloning of the polynucleotides from such genomic DNA can be effected, e.g., by using the well known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, e.g., Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligation activated transcription (LAT) and polynucleotide-based amplification (NASBA) may be used. The polynucleotides may be cloned from a strain of *Bacillus sp.*, or another or related organism from the order *Bacillales* and thus, for example, may be an allelic or species variant of the polypeptide encoding region of the polynucleotide.

30 The present invention also relates to isolated polynucleotides comprising or consisting of polynucleotides having a degree of sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 of at least 80%, e.g., 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%, at least 99%, with the proviso that it is not 100% identical to the

mature polypeptide coding sequence of SEQ ID NO: 1, and which encode a polypeptide having protease activity.

The present invention further relates to isolated polynucleotides comprising or consisting of polynucleotides having a degree of sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 3 of at least 80%, *e.g.*, 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%, at least 99%, or 100%, which encode a polypeptide having protease activity.

Modification of a polynucleotide encoding a polypeptide of the present invention may be necessary for the synthesis of polypeptides substantially similar to the polypeptide. The term "substantially similar" to the polypeptide refers to non-naturally occurring forms of the polypeptide. These polypeptides may differ in some engineered way from the polypeptide isolated from its native source, *e.g.*, variants that differ in specific activity, thermostability, pH optimum, or the like. The variant may be constructed on the basis of the polynucleotide presented as the mature polypeptide coding sequence of SEQ ID NO: 1 or SEQ ID NO: 3, *e.g.*, a subsequence thereof, and/or by introduction of nucleotide substitutions that do not result in a change in the amino acid sequence of the polypeptide, but which correspond to the codon usage of the host organism intended for production of the enzyme, or by introduction of nucleotide substitutions that may give rise to a different amino acid sequence. For a general description of nucleotide substitution, see, *e.g.*, Ford *et al.*, 1991, *Protein Expression and Purification* 2: 95-107.

The present invention also relates to isolated polynucleotides encoding polypeptides of the present invention, which hybridize under medium-high stringency conditions, high stringency conditions, or very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, (ii) the mature polypeptide coding sequence of SEQ ID NO: 3, (iii) the genomic DNA sequence comprising the mature polypeptide coding sequence of SEQ ID NO: 1, (iv) the genomic DNA sequence comprising the mature polypeptide coding sequence of SEQ ID NO: 3, or (v) the full-length complementary strand of (i), (ii), (iii), or (iv); or allelic variants and subsequences thereof (Sambrook *et al.*, 1989, *supra*), as defined herein.

In one aspect, the polynucleotide comprises or consists of SEQ ID NO: 3, the mature polypeptide coding sequence of SEQ ID NO: 3, a subsequence of SEQ ID NO: 1 that encodes a fragment of SEQ ID NO: 2, SEQ ID NO: 4 or SEQ ID NO: 5 having protease activity, or a subsequence of SEQ ID NO: 3 that encodes a fragment of SEQ ID NO: 2, SEQ ID NO: 4 or SEQ ID NO: 5 having protease activity, such as the polynucleotide of nucleotides 541 to 1089 of SEQ ID NO: 3.

Nucleic Acid Constructs

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

A polynucleotide may be manipulated in a variety of ways to provide for expression of the polypeptide. 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.

The control sequence may be a promoter sequence, a polynucleotide that is recognized by a host cell for expression of a polynucleotide encoding a polypeptide of the present invention. The promoter sequence contains transcriptional control sequences that mediate the expression of the polypeptide. The promoter may be any polynucleotide that shows transcriptional activity in the host cell of choice 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.

Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a bacterial host cell are the promoters obtained from the *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus subtilis* *xylA* and *xylB* genes, *Bacillus thuringiensis cryIIIA* gene (Agaisse and Lereclus, 1994, *Molecular Microbiology* 13: 97-107), *E. coli lac* operon, *E. coli trc* promoter (Egon *et al.*, 1988, *Gene* 69: 301-315), *Streptomyces coelicolor* agarase gene (*dagA*), and prokaryotic beta-lactamase gene (Villa-Kamaroff *et al.*, 1978, *Proc. Natl. Acad. Sci. USA* 75: 3727-3731), as well as the *tac* promoter (DeBoer *et al.*, 1983, *Proc. Natl. Acad. Sci. USA* 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in Gilbert *et al.*, 1980, *Scientific American* 242: 74-94; and in Sambrook *et al.*, 1989, *supra*. Examples of tandem promoters are disclosed in WO 99/43835.

Examples of suitable promoters for directing the 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 including a gene encoding a neutral alpha-amylase in *Aspergilli* in which the untranslated leader has been replaced by an untranslated leader from a gene encoding triose phosphate isomerase in *Aspergilli*; non-limiting examples include modified promoters including the gene encoding neutral alpha-amylase in *Aspergillus niger* in which the untranslated leader has been replaced by an untranslated leader from the gene encoding triose phosphate isomerase in *Aspergillus nidulans* or *Aspergillus oryzae*); and mutant, truncated, and hybrid promoters thereof.

In a yeast host, useful promoters are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH1, ADH2/GAP), *Saccharomyces cerevisiae* triose phosphate isomerase (TPI), *Saccharomyces cerevisiae* metallothionein (CUP1), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase. Other useful promoters for yeast host cells are described by Romanos *et al.*, 1992, *Yeast* 8: 423-488.

The control sequence may also be a suitable transcription terminator sequence, 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 polypeptide. Any terminator that is functional in the host cell of choice may be used in the present invention.

Preferred terminators for bacterial host cells are obtained from the genes for *Bacillus clausii* alkaline protease (*aprH*), *Bacillus licheniformis* alpha-amylase (*amyL*), and *Escherichia coli* ribosomal RNA (*rrnB*).

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.

Preferred terminators for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase, *Saccharomyces cerevisiae* cytochrome C (CYC1), and *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos *et al.*, 1992, *supra*.

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.

Examples of suitable mRNA stabilizer regions are obtained from a *Bacillus thuringiensis cryIIIA* gene (WO 94/25612) and a *Bacillus subtilis* SP82 gene (Hue *et al.*, 1995, *Journal of Bacteriology* 177: 3465-3471).

5 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 is operably linked to the 5'-terminus of the polynucleotide encoding the polypeptide. Any leader that is functional in the host cell may be used.

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

10 Suitable leaders for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, *Saccharomyces cerevisiae* alpha-factor, and *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP).

15 The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3'-terminus of the polynucleotide 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 of choice may be used.

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

Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, *Mol. Cellular Biol.* 15: 5983-5990.

25 The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a polypeptide and directs the polypeptide 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 polypeptide. Alternatively, the 5'-end of the coding sequence may contain a signal peptide coding sequence that is foreign to the coding sequence. The foreign signal peptide coding sequence may be required where the coding sequence does not naturally
30 contain a signal peptide coding sequence. Alternatively, the foreign signal peptide coding sequence may simply replace the natural signal peptide coding sequence in order to enhance secretion of the polypeptide. However, any signal peptide coding sequence that directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used.

35 Effective signal peptide coding sequences for bacterial host cells are the signal peptide coding sequences obtained from the genes for *Bacillus* NCIB 11837 maltogenic amylase, *Bacillus*

licheniformis subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus stearothermophilus* neutral proteases (*nprT*, *nprS*, *nprM*), and *Bacillus subtilis* *prsA*. Further signal peptides are described by Simonen and Palva, 1993, *Microbiological Reviews* 57: 109-137.

5 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 niger* glucoamylase, *Aspergillus oryzae* TAKA amylase, *Humicola insolens* cellulase, *Humicola insolens* endoglucanase V, *Humicola lanuginosa* lipase, and *Rhizomucor miehei* aspartic proteinase.

10 Useful signal peptides for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase. Other useful signal peptide coding sequences are described by Romanos *et al.*, 1992, *supra*.

 The control sequence may also be a propeptide coding sequence that encodes a propeptide positioned at the N-terminus of a polypeptide. 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 *Bacillus subtilis* alkaline protease (*aprE*), *Bacillus subtilis* neutral protease (*nprT*), *Myceliophthora thermophila* laccase (WO 95/33836), *Rhizomucor miehei* aspartic proteinase, and
15 *Saccharomyces cerevisiae* alpha-factor.
20

 Where both signal peptide and propeptide sequences are present at the N-terminus of a polypeptide, the propeptide sequence is positioned next to the N-terminus of a polypeptide and the signal peptide sequence is positioned next to the N-terminus of the propeptide sequence.

 It may also be desirable to add regulatory sequences that allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those that cause the 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 polypeptide would be operably
25 linked with the regulatory sequence.
30
35

Expression Vectors

The present invention also relates to recombinant expression vectors comprising a polynucleotide 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 (several) convenient restriction sites to allow for insertion or substitution of the polynucleotide encoding the polypeptide at such sites. Alternatively, the polynucleotide may be expressed by inserting the polynucleotide or a nucleic acid construct comprising the sequence 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.

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.

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.

The vector preferably contains one or more (several) 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.

Examples of bacterial selectable markers are *Bacillus licheniformis* or *Bacillus subtilis* *dal* genes, or markers that confer antibiotic resistance such as ampicillin, chloramphenicol, kanamycin, neomycin, spectinomycin, or tetracycline resistance. Suitable markers for yeast host cells include, but are not limited to, ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. 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 adenylyltransferase), and *trpC* (anthranilate synthase), as well as equivalents thereof.

Preferred for use in an *Aspergillus* cell are *Aspergillus nidulans* or *Aspergillus oryzae amdS* and *pyrG* genes and a *Streptomyces hygrosopicus bar* gene.

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.

5 For integration into the host cell genome, the vector may rely on the polynucleotide's sequence encoding the polypeptide 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
10 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,
15 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.

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
20 "origin of replication" or "plasmid replicator" means a polynucleotide that enables a plasmid or vector to replicate *in vivo*.

Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAM β 1 permitting replication in *Bacillus*.

25 Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6.

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;
30 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.

More than one copy of a polynucleotide of the present invention may be inserted into a host cell to increase production of a polypeptide. 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
35 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.

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., 5 Sambrook *et al.*, 1989, *supra*).

Host Cells

The present invention also relates to recombinant host cells, comprising a polynucleotide of the present invention operably linked to one or more (several) control sequences that direct the 10 production of a polypeptide of the present invention. 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 15 upon the gene encoding the polypeptide and its source.

The host cell may be any cell useful in the recombinant production of a polypeptide of the present invention, e.g., a prokaryote or a eukaryote.

The prokaryotic host cell may be any Gram-positive or Gram-negative bacterium. Gram-positive bacteria include, but are not limited to, *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, 20 *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, and *Streptomyces*. Gram-negative bacteria include, but are not limited to, *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, and *Ureaplasma*.

The bacterial host cell may be any *Bacillus* cell including, but not limited to, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus* 25 *coagulans*, *Bacillus firmus*, *Geobacillus stearothermophilus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis* cells. Specifically preferred host cells are *Bacillus subtilis* and *Bacillus licheniformis* cells.

The bacterial host cell may also be any *Streptococcus* cell including, but not limited to, 30 *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, and *Streptococcus equi* subsp. *Zooepidemicus* cells.

The bacterial host cell may also be any *Streptomyces* cell including, but not limited to, *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, and *Streptomyces lividans* cells.

The introduction of DNA into a *Bacillus* cell may, for instance, be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, *Mol. Gen. Genet.* 168: 111-115), by using competent cells (see, e.g., Young and Spizizen, 1961, *J. Bacteriol.* 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, *J. Mol. Biol.* 56: 209-221), by electroporation (see, e.g., Shigekawa and Dower, 1988, *Biotechniques* 6: 742-751), or by conjugation (see, e.g., Koehler and Thorne, 1987, *J. Bacteriol.* 169: 5271-5278). The introduction of DNA into an *E. coli* cell may, for instance, be effected by protoplast transformation (see, e.g., Hanahan, 1983, *J. Mol. Biol.* 166: 557-580) or electroporation (see, e.g., Dower *et al.*, 1988, *Nucleic Acids Res.* 16: 6127-6145). The introduction of DNA into a *Streptomyces* cell may, for instance, be effected by protoplast transformation and electroporation (see, e.g., Gong *et al.*, 2004, *Folia Microbiol. (Praha)* 49: 399-405), by conjugation (see, e.g., Mazodier *et al.*, 1989, *J. Bacteriol.* 171: 3583-3585), or by transduction (see, e.g., Burke *et al.*, 2001, *Proc. Natl. Acad. Sci. USA* 98: 6289-6294). The introduction of DNA into a *Pseudomonas* cell may, for instance, be effected by electroporation (see, e.g., Choi *et al.*, 2006, *J. Microbiol. Methods* 64: 391-397) or by conjugation (see, e.g., Pinedo and Smets, 2005, *Appl. Environ. Microbiol.* 71: 51-57). The introduction of DNA into a *Streptococcus* cell may, for instance, be effected by natural competence (see, e.g., Perry and Kuramitsu, 1981, *Infect. Immun.* 32: 1295-1297), by protoplast transformation (see, e.g., Catt and Jollick, 1991, *Microbios* 68: 189-207), by electroporation (see, e.g., Buckley *et al.*, 1999, *Appl. Environ. Microbiol.* 65: 3800-3804) or by conjugation (see, e.g., Clewell, 1981, *Microbiol. Rev.* 45: 409-436). However, any method known in the art for introducing DNA into a host cell can be used.

The host cell may also be a eukaryote, such as a mammalian, insect, plant, or fungal cell.

The host cell may be a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota (as defined by Hawksworth *et al.*, *In, Ainsworth and Bisby's Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK) as well as the Oomycota (as cited in Hawksworth *et al.*, 1995, *supra*, page 171) and all mitosporic fungi (Hawksworth *et al.*, 1995, *supra*).

The fungal host cell may be a yeast cell. "Yeast" as used herein includes ascosporeogenous yeast (Endomycetales), basidiosporeogenous 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, F.A., Passmore, S.M., and Davenport, R.R., eds, *Soc. App. Bacteriol. Symposium Series No. 9*, 1980).

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.

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.

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*, *Tolyptocladium*, *Trametes*, or *Trichoderma* cell.

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.

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 and Yelton *et al.*, 1984, *Proc. Natl. Acad. Sci. USA* 81: 1470-1474. 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.

5 **Methods of Production**

The present invention also relates to methods of producing a polypeptide of the present invention, comprising: (a) cultivating a cell, which in its wild-type form produces the polypeptide, under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide. In a preferred aspect, the cell is of the genus *Bacillus*. In a more preferred aspect, the cell is *Bacillus* sp-19138.

The present invention also relates to methods of producing a polypeptide of the present invention, comprising: (a) cultivating a recombinant host cell of the present invention under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide.

The host cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods well known in the art. For example, the cell may be cultivated by shake flask cultivation, and 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 polypeptide 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 polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

More details are provided in the Section on “Nucleic Acid Constructs, Expression Vectors, Recombinant Host Cells, and Methods for Production of Proteases” below.

The polypeptide may be detected using methods known in the art that are specific for the polypeptides. These detection methods may include 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 polypeptide.

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

The polypeptide 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*, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989) to obtain substantially pure polypeptides.

In an alternative aspect, the polypeptide is not recovered, but rather a host cell of the present invention expressing a polypeptide is used as a source of the polypeptide.

Plants

The present invention also relates to plants, e.g., a transgenic plant, plant part, or plant cell, comprising an isolated polynucleotide of the present invention so as to express and produce the polypeptide in recoverable quantities. The polypeptide may be recovered from the plant or plant part. Alternatively, the plant or plant part containing the polypeptide may be used as such for improving the quality of a food or feed, e.g., improving nutritional value, palatability, and rheological properties, or to destroy an antinutritive factor.

The transgenic plant can be dicotyledonous (a dicot) or monocotyledonous (a monocot). Examples of monocot plants are grasses, such as meadow grass (blue grass, *Poa*), forage grass such as *Festuca*, *Lolium*, temperate grass, such as *Agrostis*, and cereals, e.g., wheat, oats, rye, barley, rice, sorghum, and maize (corn).

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

Examples of plant parts are stem, callus, leaves, root, fruits, seeds, and tubers as well as the individual tissues comprising these parts, e.g., epidermis, mesophyll, parenchyme, vascular tissues, meristems. Specific plant cell compartments, such as chloroplasts, apoplasts, mitochondria, vacuoles, peroxisomes and cytoplasm are also considered to be a plant part. Furthermore, any plant cell, whatever the tissue origin, is considered to be a plant part. Likewise, plant parts such as specific tissues and cells isolated to facilitate the utilization of the invention are also considered plant parts, e.g., embryos, endosperms, aleurone and seeds coats.

Also included within the scope of the present invention are the progeny of such plants, plant parts, and plant cells.

The transgenic plant or plant cell expressing a polypeptide may be constructed in accordance with methods known in the art. In short, the plant or plant cell is constructed by incorporating one or more (several) expression constructs encoding a polypeptide into the plant

host genome or chloroplast genome and propagating the resulting modified plant or plant cell into a transgenic plant or plant cell.

The expression construct is conveniently a nucleic acid construct that comprises a polynucleotide encoding a polypeptide operably linked with appropriate regulatory sequences required for expression of the polynucleotide in the plant or plant part of choice. Furthermore, the expression construct may comprise a selectable marker useful for identifying host cells into which the expression construct has been integrated and DNA sequences necessary for introduction of the construct into the plant in question (the latter depends on the DNA introduction method to be used).

The choice of regulatory sequences, such as promoter and terminator sequences and optionally signal or transit sequences, is determined, for example, on the basis of when, where, and how the polypeptide is desired to be expressed. For instance, the expression of the gene encoding a polypeptide may be constitutive or inducible, or may be developmental, stage or tissue specific, and the gene product may be targeted to a specific tissue or plant part such as seeds or leaves. Regulatory sequences are, for example, described by Tague *et al.*, 1988, *Plant Physiology* 86: 506.

For constitutive expression, the 35S-CaMV, the maize ubiquitin 1, and the rice actin 1 promoter may be used (Franck *et al.*, 1980, *Cell* 21: 285-294; Christensen *et al.*, 1992, *Plant Mol. Biol.* 18: 675-689; Zhang *et al.*, 1991, *Plant Cell* 3: 1155-1165). Organ-specific promoters may be, for example, a promoter from storage sink tissues such as seeds, potato tubers, and fruits (Edwards and Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303), or from metabolic sink tissues such as meristems (Ito *et al.*, 1994, *Plant Mol. Biol.* 24: 863-878), a seed specific promoter such as the glutelin, prolamin, globulin, or albumin promoter from rice (Wu *et al.*, 1998, *Plant Cell Physiol.* 39: 885-889), a *Vicia faba* promoter from the legumin B4 and the unknown seed protein gene from *Vicia faba* (Conrad *et al.*, 1998, *J. Plant Physiol.* 152: 708-711), a promoter from a seed oil body protein (Chen *et al.*, 1998, *Plant Cell Physiol.* 39: 935-941), the storage protein *napA* promoter from *Brassica napus*, or any other seed specific promoter known in the art, *e.g.*, as described in WO 91/14772. Furthermore, the promoter may be a leaf specific promoter such as the *rbcs* promoter from rice or tomato (Kyoizuka *et al.*, 1993, *Plant Physiol.* 102: 991-1000), the chlorella virus adenine methyltransferase gene promoter (Mitra and Higgins, 1994, *Plant Mol. Biol.* 26: 85-93), the *aldP* gene promoter from rice (Kagaya *et al.*, 1995, *Mol. Gen. Genet.* 248: 668-674), or a wound inducible promoter such as the potato *pin2* promoter (Xu *et al.*, 1993, *Plant Mol. Biol.* 22: 573-588). Likewise, the promoter may be inducible by abiotic treatments such as temperature, drought, or alterations in salinity or induced by exogenously applied substances that activate the promoter, *e.g.*, ethanol, oestrogens, plant hormones such as ethylene, abscisic acid, and gibberellic acid, and heavy metals.

A promoter enhancer element may also be used to achieve higher expression of a polypeptide in the plant. For instance, the promoter enhancer element may be an intron that is placed between the promoter and the polynucleotide encoding a polypeptide. For instance, Xu *et al.*, 1993, *supra*, disclose the use of the first intron of the rice actin 1 gene to enhance expression.

5 The selectable marker gene and any other parts of the expression construct may be chosen from those available in the art.

The nucleic acid construct is incorporated into the plant genome according to conventional techniques known in the art, including *Agrobacterium*-mediated transformation, virus-mediated transformation, microinjection, particle bombardment, biolistic transformation, and electroporation
10 (Gasser *et al.*, 1990, *Science* 244: 1293; Potrykus, 1990, *Bio/Technology* 8: 535; Shimamoto *et al.*, 1989, *Nature* 338: 274).

Presently, *Agrobacterium tumefaciens*-mediated gene transfer is the method of choice for generating transgenic dicots (for a review, see Hooykas and Schilperoort, 1992, *Plant Mol. Biol.* 19: 15-38) and can also be used for transforming monocots, although other transformation methods are
15 often used for these plants. Presently, the method of choice for generating transgenic monocots is particle bombardment (microscopic gold or tungsten particles coated with the transforming DNA) of embryonic calli or developing embryos (Christou, 1992, *Plant J.* 2: 275-281; Shimamoto, 1994, *Curr. Opin. Biotechnol.* 5: 158-162; Vasil *et al.*, 1992, *Bio/Technology* 10: 667-674). An alternative method for transformation of monocots is based on protoplast transformation as described by
20 Omirulleh *et al.*, 1993, *Plant Mol. Biol.* 21: 415-428. Additional transformation methods for use in accordance with the present disclosure include those described in U.S. Patent Nos. 6,395,966 and 7,151,204 (both of which are herein incorporated by reference in their entirety).

Following transformation, the transformants having incorporated the expression construct are selected and regenerated into whole plants according to methods well known in the art. Often
25 the transformation procedure is designed for the selective elimination of selection genes either during regeneration or in the following generations by using, for example, co-transformation with two separate T-DNA constructs or site specific excision of the selection gene by a specific recombinase.

In addition to direct transformation of a particular plant genotype with a construct prepared
30 according to the present invention, transgenic plants may be made by crossing a plant having the construct to a second plant lacking the construct. For example, a construct encoding a polypeptide can be introduced into a particular plant variety by crossing, without the need for ever directly transforming a plant of that given variety. Therefore, the present invention encompasses not only a plant directly regenerated from cells which have been transformed in accordance with the present
35 invention, but also the progeny of such plants. As used herein, progeny may refer to the offspring of

any generation of a parent plant prepared in accordance with the present invention. Such progeny may include a DNA construct prepared in accordance with the present invention, or a portion of a DNA construct prepared in accordance with the present invention. Crossing results in the introduction of a transgene into a plant line by cross pollinating a starting line with a donor plant
5 line. Non-limiting examples of such steps are further articulated in U.S. Patent No: 7,151,204.

Plants may be generated through a process of backcross conversion. For example, plants include plants referred to as a backcross converted genotype, line, inbred, or hybrid.

Genetic markers may be used to assist in the introgression of one or more transgenes of the invention from one genetic background into another. Marker assisted selection offers
10 advantages relative to conventional breeding in that it can be used to avoid errors caused by phenotypic variations. Further, genetic markers may provide data regarding the relative degree of elite germplasm in the individual progeny of a particular cross. For example, when a plant with a desired trait which otherwise has a non-agronomically desirable genetic background is crossed to an elite parent, genetic markers may be used to select progeny which not only possess the trait of
15 interest, but also have a relatively large proportion of the desired germplasm. In this way, the number of generations required to introgress one or more traits into a particular genetic background is minimized.

The present invention also relates to methods of producing a polypeptide of the present invention comprising: (a) cultivating a transgenic plant or a plant cell comprising a polynucleotide
20 encoding the polypeptide under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide.

Compositions

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

The composition may comprise a protease of the present invention as the major enzymatic component, *e.g.*, a mono-component composition. Alternatively, the composition may comprise
30 multiple enzymatic activities, such as an aminopeptidase, amylase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, alpha-galactosidase, beta-galactosidase, glucoamylase, alpha-glucosidase, beta-glucosidase, haloperoxidase, invertase, laccase, lipase, mannosidase, oxidase, pectinolytic enzyme, peptidoglutaminase, peroxidase, phytase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transglutaminase, or xylanase. The additional enzyme(s) may be produced,

for example, by microorganisms such as bacteria or fungi or by plants or by animals. 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 a granulate or a microgranulate. The protease may be stabilized in accordance with methods known in the art.

Uses

The present invention is also directed to methods for using the polypeptides having protease activity, or compositions thereof, for e.g. animal feed.

Animal Feed

The present invention is also directed to methods for using the proteases having protease activity in animal feed, as well as to feed compositions and feed additives comprising the proteases of the invention.

The term animal includes all animals, including human beings. Examples of animals are non-ruminants, and ruminants. Ruminant animals include, for example, animals such as sheep, goats, and cattle, e.g. beef cattle, cows, and young calves. In a particular embodiment, the animal is a non-ruminant animal. Non-ruminant animals include mono-gastric animals, e.g. pigs or swine (including, but not limited to, piglets, growing pigs, and sows); poultry such as turkeys, ducks and chicken (including but not limited to broiler chicks, layers); horses (including but not limited to hotbloods, coldbloods and warm bloods), young calves; and fish (including but not limited to salmon, trout, tilapia, catfish and carps; and crustaceans (including but not limited to shrimps and prawns).

The term feed or feed composition means any compound, preparation, mixture, or composition suitable for, or intended for intake by an animal.

In the use according to the invention the protease can be fed to the animal before, after, or simultaneously with the diet. The latter is preferred.

In a particular embodiment, the protease, in the form in which it is added to the feed, or when being included in a feed additive, is well-defined. Well-defined means that the protease preparation is at least 50% pure as determined by Size-exclusion chromatography (see Example 12 of WO 01/58275). In other particular embodiments the protease preparation is at least 60, 70, 80, 85, 88, 90, 92, 94, or at least 95% pure as determined by this method.

A well-defined protease preparation is advantageous. For instance, it is much easier to dose correctly to the feed a protease that is essentially free from interfering or contaminating other

proteases. The term dose correctly refers in particular to the objective of obtaining consistent and constant results, and the capability of optimising dosage based upon the desired effect.

For the use in animal feed, however, the protease need not be that pure; it may e.g. include other enzymes, in which case it could be termed a protease preparation.

5 The protease preparation can be (a) added directly to the feed (or used directly in a protein treatment process), or (b) it can be used in the production of one or more intermediate compositions such as feed additives or premixes that is subsequently added to the feed (or used in a treatment process). The degree of purity described above refers to the purity of the original protease preparation, whether used according to (a) or (b) above.

10 Protease preparations with purities of this order of magnitude are in particular obtainable using recombinant methods of production, whereas they are not so easily obtained and also subject to a much higher batch-to-batch variation when the protease is produced by traditional fermentation methods.

Such protease preparation may of course be mixed with other enzymes.

15 The protein may be an animal protein, such as meat and bone meal, feather meal, and/or fish meal; or it may be a vegetable protein.

20 The term vegetable proteins as used herein refers to any compound, composition, preparation or mixture that includes at least one protein derived from or originating from a vegetable, including modified proteins and protein-derivatives. In particular embodiments, the protein content of the vegetable proteins is at least 10, 20, 30, 40, 50, or 60% (w/w).

Vegetable proteins may be derived from vegetable protein sources, such as legumes and cereals, for example materials from plants of the families *Fabaceae* (*Leguminosae*), *Cruciferaeae*, *Chenopodiaceae*, and *Poaceae*, such as soy bean meal, lupin meal and rapeseed meal.

25 In a particular embodiment, the vegetable protein source is material from one or more plants of the family *Fabaceae*, e.g. soybean, lupine, pea, or bean.

In another particular embodiment, the vegetable protein source is material from one or more plants of the family *Chenopodiaceae*, e.g. beet, sugar beet, spinach or quinoa.

Other examples of vegetable protein sources are rapeseed, sunflower seed, cotton seed, and cabbage.

30 Soybean is a preferred vegetable protein source.

Other examples of vegetable protein sources are cereals such as barley, wheat, rye, oat, maize (corn), rice, triticale, and sorghum.

35 In a particular embodiment of a treatment process the protease(s) in question is affecting (or acting on, or exerting its hydrolyzing or degrading influence on) the proteins, such as vegetable proteins or protein sources. To achieve this, the protein or protein source is typically suspended in

a solvent, eg an aqueous solvent such as water, and the pH and temperature values are adjusted paying due regard to the characteristics of the enzyme in question. For example, the treatment may take place at a pH-value at which the activity of the actual protease is at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or at least 90%. Likewise, for example, the treatment may take place
5 at a temperature at which the activity of the actual protease is at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or at least 90%. The above percentage activity indications are relative to the maximum activities. The enzymatic reaction is continued until the desired result is achieved, following which it may or may not be stopped by inactivating the enzyme, e.g. by a heat-treatment step.

10 In another particular embodiment of a treatment process of the invention, the protease action is sustained, meaning e.g. that the protease is added to the proteins, but its hydrolysing influence is so to speak not switched on until later when desired, once suitable hydrolysing conditions are established, or once any enzyme inhibitors are inactivated, or whatever other means could have been applied to postpone the action of the enzyme.

15 In one embodiment the treatment is a pre-treatment of animal feed or proteins for use in animal feed, i.e. the proteins are hydrolysed before intake.

The term improving the nutritional value of an animal feed means improving the availability of nutrients in the feed. In this invention improving the nutritional values refers in particular to improving the availability of the protein fraction of the feed, thereby leading to increased protein
20 extraction, higher protein yields, and/or improved protein utilization. When the nutritional value of the feed is increased, the protein and/or amino acid digestibility is increased and the growth rate and/or weight gain and/or feed conversion (i.e. the weight of ingested feed relative to weight gain) of the animal might be improved.

The protease can be added to the feed in any form, be it as a relatively pure protease or in
25 admixture with other components intended for addition to animal feed, i.e. in the form of animal feed additives, such as the so-called pre-mixes for animal feed.

In a further aspect the present invention relates to compositions for use in animal feed, such as animal feed, and animal feed additives, e.g. premixes.

30 Apart from the protease of the invention, the animal feed additives of the invention contain at least one fat-soluble vitamin, and/or at least one water soluble vitamin, and/or at least one trace mineral, and/or at least one macro mineral.

Further, optional, feed-additive ingredients are colouring agents, e.g. carotenoids such as beta-carotene, astaxanthin, and lutein; stabilisers; growth improving additives and aroma compounds/flavorings, e.g. creosol, anethol, deca-, undeca-and/or dodeca-lactones, ionones,
35 irone, gingerol, piperidine, propylidene phthalide, butylidene phthalide, capsaicin and/or tannin;

antimicrobial peptides; polyunsaturated fatty acids (PUFAs); reactive oxygen generating species; also, a support may be used that may contain, for example, 40-50% by weight of wood fibres, 8-10% by weight of stearine, 4-5% by weight of curcuma powder, 4-58% by weight of rosemary powder, 22-28% by weight of limestone, 1-3% by weight of a gum, such as gum arabic, 5-50% by weight of sugar and/or starch and 5-15% by weight of water.

A feed or a feed additive of the invention may also comprise at least one other enzyme selected from amongst phytase (EC 3.1.3.8 or 3.1.3.26); xylanase (EC 3.2.1.8); galactanase (EC 3.2.1.89); alpha-galactosidase (EC 3.2.1.22); further protease (EC 3.4), phospholipase A1 (EC 3.1.1.32); phospholipase A2 (EC 3.1.1.4); lysophospholipase (EC 3.1.1.5); phospholipase C (3.1.4.3); phospholipase D (EC 3.1.4.4); amylase such as, for example, alpha-amylase (EC 3.2.1.1); and/or beta-glucanase (EC 3.2.1.4 or EC 3.2.1.6).

In a particular embodiment these other enzymes are well-defined (as defined above for protease preparations).

Examples of antimicrobial peptides (AMP's) are CAP18, Leucocin A, Tritrpticin, Protegrin-1, Thanatin, Defensin, Lactoferrin, Lactoferricin, and Ovispirin such as Novispirin (Robert Lehrer, 2000), Plectasins, and Statins, including the compounds and polypeptides disclosed in WO 03/044049 and WO 03/048148, as well as variants or fragments of the above that retain antimicrobial activity.

Examples of antifungal polypeptides (AFP's) are the *Aspergillus giganteus*, and *Aspergillus niger* peptides, as well as variants and fragments thereof which retain antifungal activity, as disclosed in WO 94/01459 and WO 02/090384.

Examples of polyunsaturated fatty acids are C18, C20 and C22 polyunsaturated fatty acids, such as arachidonic acid, docosohexaenoic acid, eicosapentaenoic acid and gamma-linoleic acid.

Examples of reactive oxygen generating species are chemicals such as perborate, persulphate, or percarbonate; and enzymes such as an oxidase, an oxygenase or a syntethase.

Usally fat- and water-soluble vitamins, as well as trace minerals form part of a so-called premix intended for addition to the feed, whereas macro minerals are usually separately added to the feed. Either of these composition types, when enriched with a protease of the invention, is an animal feed additive of the invention.

In a particular embodiment, the animal feed additive of the invention is intended for being included (or prescribed as having to be included) in animal diets or feed at levels of 0.01 to 10.0%; more particularly 0.05 to 5.0%; or 0.2 to 1.0% (% meaning g additive per 100 g feed). This is so in particular for premixes.

The following are non-exclusive lists of examples of these components:

Examples of fat-soluble vitamins are vitamin A, vitamin D3, vitamin E, and vitamin K, e.g. vitamin K3.

Examples of water-soluble vitamins are vitamin B12, biotin and choline, vitamin B1, vitamin B2, vitamin B6, niacin, folic acid and panthothenate, e.g. Ca-D-panthothenate.

5 Examples of trace minerals are manganese, zinc, iron, copper, iodine, selenium, and cobalt.
 Examples of macro minerals are calcium, phosphorus and sodium.

The nutritional requirements of these components (exemplified with poultry and piglets/pigs) are listed in Table A of WO 01/58275. Nutritional requirement means that these components should be provided in the diet in the concentrations indicated.

10 In the alternative, the animal feed additive of the invention comprises at least one of the individual components specified in Table A of WO 01/58275. At least one means either of, one or more of, one, or two, or three, or four and so forth up to all thirteen, or up to all fifteen individual components. More specifically, this at least one individual component is included in the additive of the invention in such an amount as to provide an in-feed-concentration within the range indicated in
 15 column four, or column five, or column six of Table A.

In a still further embodiment, the animal feed additive of the invention comprises at least one of the below vitamins, preferably to provide an in-feed-concentration within the ranges specified in the below Table 1 (for piglet diets, and broiler diets, respectively).

Table 1: Typical vitamin recommendations

Vitamin	Piglet diet	Broiler diet
Vitamin A	10,000-15,000 IU/kg feed	8-12,500 IU/kg feed
Vitamin D3	1800-2000 IU/kg feed	3000-5000 IU/kg feed
Vitamin E	60-100 mg/kg feed	150-240 mg/kg feed
Vitamin K3	2-4 mg/kg feed	2-4 mg/kg feed
Vitamin B1	2-4 mg/kg feed	2-3 mg/kg feed
Vitamin B2	6-10 mg/kg feed	7-9 mg/kg feed
Vitamin B6	4-8 mg/kg feed	3-6 mg/kg feed
Vitamin B12	0.03-0.05 mg/kg feed	0.015-0.04 mg/kg feed
Niacin (Vitamin B3)	30-50 mg/kg feed	50-80 mg/kg feed
Pantothenic acid	20-40 mg/kg feed	10-18 mg/kg feed

Folic acid	1-2 mg/kg feed	1-2 mg/kg feed
Biotin	0.15-0.4 mg/kg feed	0.15-0.3 mg/kg feed
Choline chloride	200-400 mg/kg feed	300-600 mg/kg feed

The present invention also relates to animal feed compositions. Animal feed compositions or diets have a relatively high content of protein. Poultry and pig diets can be characterised as indicated in Table B of WO 01/58275, columns 2-3. Fish diets can be characterised as indicated in column 4 of this Table B. Furthermore such fish diets usually have a crude fat content of 200-310 g/kg.

WO 01/58275 corresponds to US 09/779334 which is hereby incorporated by reference.

An animal feed composition according to the invention has a crude protein content of 50-800 g/kg, and furthermore comprises at least one protease as claimed herein.

Furthermore, or in the alternative (to the crude protein content indicated above), the animal feed composition of the invention has a content of metabolisable energy of 10-30 MJ/kg; and/or a content of calcium of 0.1-200 g/kg; and/or a content of available phosphorus of 0.1-200 g/kg; and/or a content of methionine of 0.1-100 g/kg; and/or a content of methionine plus cysteine of 0.1-150 g/kg; and/or a content of lysine of 0.5-50 g/kg.

In particular embodiments, the content of metabolisable energy, crude protein, calcium, phosphorus, methionine, methionine plus cysteine, and/or lysine is within any one of ranges 2, 3, 4 or 5 in Table B of WO 01/58275 (R. 2-5).

Crude protein is calculated as nitrogen (N) multiplied by a factor 6.25, i.e. Crude protein (g/kg)= N (g/kg) x 6.25. The nitrogen content is determined by the Kjeldahl method (A.O.A.C., 1984, Official Methods of Analysis 14th ed., Association of Official Analytical Chemists, Washington DC).

Metabolisable energy can be calculated on the basis of the NRC publication Nutrient requirements in swine, ninth revised edition 1988, subcommittee on swine nutrition, committee on animal nutrition, board of agriculture, national research council. National Academy Press, Washington, D.C., pp. 2-6, and the European Table of Energy Values for Poultry Feed-stuffs, Spelderholt centre for poultry research and extension, 7361 DA Beekbergen, The Netherlands. Grafisch bedrijf Ponsen & looijen bv, Wageningen. ISBN 90-71463-12-5.

The dietary content of calcium, available phosphorus and amino acids in complete animal diets is calculated on the basis of feed tables such as Veevoedertabel 1997, gegevens over chemische samenstelling, verteerbaarheid en voederwaarde van voedermiddelen, Central Veevoederbureau, Runderweg 6, 8219 pk Lelystad. ISBN 90-72839-13-7.

In a particular embodiment, the animal feed composition of the invention contains at least one vegetable protein as defined above.

The animal feed composition of the invention may also contain animal protein, such as Meat and Bone Meal, Feather meal, and/or Fish Meal, typically in an amount of 0-25%. The animal feed composition of the invention may also comprise Dried Destillers Grains with Solubles (DDGS),
5 typically in amounts of 0-30%.

In still further particular embodiments, the animal feed composition of the invention contains 0-80% maize; and/or 0-80% sorghum; and/or 0-70% wheat; and/or 0-70% Barley; and/or 0-30%
10 oats; and/or 0-40% soybean meal; and/or 0-25% fish meal; and/or 0-25% meat and bone meal; and/or 0-20% whey.

Animal diets can e.g. be manufactured as mash feed (non pelleted) or pelleted feed. Typically, the milled feed-stuffs are mixed and sufficient amounts of essential vitamins and minerals are added according to the specifications for the species in question. Enzymes can be added as solid or liquid enzyme formulations. For example, for mash feed a solid or liquid enzyme
15 formulation may be added before or during the ingredient mixing step. For pelleted feed the (liquid or solid) protease/enzyme preparation may also be added before or during the feed ingredient step. Typically a liquid protease/enzyme preparation is added after the pelleting step. The enzyme may also be incorporated in a feed additive or premix.

The final enzyme concentration in the diet is within the range of 0.01-200 mg enzyme
20 protein per kg diet, for example in the range of 0.5-25 mg enzyme protein per kg animal diet.

The protease should of course be applied in an effective amount, i.e. in an amount adequate for improving hydrolysis, digestibility, and/or improving nutritional value of feed. It is at present contemplated that the enzyme is administered in one or more of the following amounts (dosage ranges): 0.01-200; 0.01-100; 0.5-100; 1-50; 5-100; 10-100; 0.05-50; or 0.10-10 – all these
25 ranges being in mg protease protein per kg feed (ppm).

For determining mg protease protein per kg feed, the protease is purified from the feed composition, and the specific activity of the purified protease is determined using a relevant assay (see under protease activity, substrates, and assays). The protease activity of the feed composition as such is also determined using the same assay, and on the basis of these two determinations,
30 the dosage in mg protease protein per kg feed is calculated.

The same principles apply for determining mg protease protein in feed additives. Of course, if a sample is available of the protease used for preparing the feed additive or the feed, the specific activity is determined from this sample (no need to purify the protease from the feed composition or the additive).

Nucleic Acid Constructs, Expression Vectors, Recombinant Host Cells, and Methods for Production of Proteases

The present invention also relates to nucleic acid constructs, expression vectors and recombinant host cells comprising such polynucleotides encoding the proteases of the invention.

5 The present invention also relates to methods of producing a protease, comprising: (a) cultivating a recombinant host cell comprising such polynucleotide; and (b) recovering the protein.

The protein may be native or heterologous to a host cell. The term "protein" is not meant herein to refer to a specific length of the encoded product and, therefore, encompasses peptides, oligopeptides, and proteins. The term "protein" also encompasses two or more polypeptides
10 combined to form the encoded product. The proteins also include hybrid polypeptides and fused polypeptides.

Preferably, the protein is a protease. For example, the protein may be a hydrolase, such as a proteolytic enzyme or protease. The gene may be obtained from any prokaryotic, eukaryotic, or
15 other source.

The present invention is further described by the following examples that should not be construed as limiting the scope of the invention.

EXAMPLES

Materials and Methods

20 Protease assays

1) Suc-AAPF-pNA assay

pNA substrate : Suc-AAPF-pNA (Bachem L-1400).

Temperature : Room temperature (25°C)

25 Assay buffers : 100mM succinic acid, 100mM HEPES, 100mM CHES, 100mM CABS,
1mM CaCl₂, 150mM KCl, 0.01% Triton X-100 adjusted to pH-values 2.0,
3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, and 11.0 with HCl or NaOH.

20µl protease (diluted in 0.01% Triton X-100) was mixed with 100µl assay buffer. The assay was started by adding 100µl pNA substrate (50mg dissolved in 1.0ml DMSO and further diluted 45x
30 with 0.01% Triton X-100). The increase in OD₄₀₅ was monitored as a measure of the protease activity.

2) Protazyme AK assay

Substrate : Protazyme AK tablet (cross-linked and dyed casein; from Megazyme)

Temperature : controlled (assay temperature).

Assay buffers : 100mM succinic acid, 100mM HEPES, 100mM CHES, 100mM CABS,
 5 1mM CaCl₂, 150mM KCl, 0.01% Triton X-100 adjusted to pH-values 2.0,
 3.0, 4.0, 5.0, 6.0, 6.5, 7.0, 8.0, 9.0, 10.0, and 11.0 with HCl or NaOH.

A Protazyme AK tablet was suspended in 2.0ml 0.01% Triton X-100 by gentle stirring. 500µl
 of this suspension and 500µl assay buffer were dispensed in an Eppendorf tube and placed on ice.
 10 20µl protease sample (diluted in 0.01% Triton X-100) was added. The assay was initiated by
 transferring the Eppendorf tube to an Eppendorf thermomixer, which was set to the assay
 temperature. The tube was incubated for 15 minutes on the Eppendorf thermomixer at its highest
 shaking rate (1400 rpm.). The incubation was stopped by transferring the tube back to the ice bath.
 Then the tube was centrifuged in an ice cold centrifuge for a few minutes and 200µl supernatant
 15 was transferred to a microtiter plate. OD₆₅₀ was read as a measure of protease activity. A buffer
 blind was included in the assay (instead of enzyme).

3) Suc-AAPX-pNA assay

pNA substrates: Suc-AAPA-pNA (Bachem L-1775)

20 Suc-AAPR-pNA (Bachem L-1720)

Suc-AAPD-pNA (Bachem L-1835)

Suc-AAPI-pNA (Bachem L-1790)

Suc-AAPM-pNA (Bachem L-1395)

Suc-AAPV-pNA (Bachem L-1770)

25 Suc-AAPL-pNA (Bachem L-1390)

Suc-AAPE-pNA (Bachem L-1710)

Suc-AAPK-pNA (Bachem L-1725)

Suc-AAPF-pNA (Bachem L-1400)

Temperature : Room temperature (25°C)

30 Assay buffer : 100mM succinic acid, 100mM HEPES, 100mM CHES, 100mM CABS,
 1mM CaCl₂, 150mM KCl, 0.01% Triton X-100, pH 6.0 or pH 9.0.

20µl protease (diluted in 0.01% Triton X-100) was mixed with 100µl assay buffer. The assay
 was started by adding 100µl pNA substrate (50mg dissolved in 1.0ml DMSO and further diluted 45x

with 0.01% Triton X-100). The increase in OD₄₀₅ was monitored as a measure of the protease activity.

o-Phthaldialdehyde (OPA) assay

This assay detects primary amines and hence cleavage of peptide bonds by a protease can be measured as the difference in absorbance between a protease treated sample and a control sample. The assay is conducted essentially according to Nielsen *et al.* (Nielsen, PM, Petersen, D, Dampmann, C. Improved method for determining food protein degree of hydrolysis, 2001, *J Food Sci*, 66: 642-646).

500 µl of sample is filtered through a 100 kDa Microcon centrifugal filter (60 min, 11,000 rpm, 5°C). The samples are diluted appropriately (e.g. 10, 50 or 100 times) in deionizer water and 25 µl of each sample is loaded into a 96 well microtiter plate (5 replicates). 200 µl OPA reagent (100 mM di-sodium tetraborate decahydrate, 3.5 mM sodium dodecyl sulphate (SDS), 5.7 mM di-thiothreitol (DDT), 6 mM o-phthaldialdehyde) is dispensed into all wells, the plate is shaken (10 sec, 750 rpm) and absorbance measured at 340 nm.

15 **Strain**

The nucleotide sequence encoding the S1 protease 1 from *Saccharopolyspora erythraea* was published by Oliynyk et al in 'Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338', 2007, *Nat. Biotechnol.* 25:447-453 and the gene was submitted to EMBL/GenBank under accession number EMBL: AM420293. According to Oliynyk, 'the strain used, NRRL23338, is the original form of the type strain of *S. erythraea* NRRL23338, which is now listed as NRRL23338 white in the NRRL database'. The NRRL database indicates (under NRRL number B-24071 which corresponds to NRRL 23338 white) that the 'white colony variant was isolated from growth from ampule from second lyophilization'. The reference for NRRL23338 refers to US 2,653,899 wherein it is stated that the original sample of *S. erythraea* NRRL23338 was obtained as a soil sample from Ilonio City, Phillipine Islands on or before 1952.

Example 1: Expression of the S1 protease 2 from *Saccharopolyspora erythraea*

Based on the published nucleotide sequence identified as SEQ ID NO: 1, a synthetic gene having SEQ ID NO: 3 was synthesized by Gene Art (GENEART AG BioPark, Josef-Engert-Str. 11, 93053, Regensburg, Germany). The synthetic gene was subcloned using ClaI and MluI restriction sites into a *Bacillus* expression vector as described in WO 12/025577. Transformants were selected on LB plates supplemented with 6 µg of chloramphenicol per ml. The recombinant *Bacillus subtilis* clone containing the integrated expression construct was selected and designated as S1

protease 2 from *Saccharopolyspora erythraea*. It was cultivated on a rotary shaking table in 500 mL baffled Erlenmeyer flasks each containing 100 ml casein-based media supplemented with 34 mg/l chloramphenicol. The clone was cultivated for 5 days at 26°C, at 225 rpm. The enzyme containing supernatants were harvested and the enzyme purified as described in Example 2.

5 **Example 2: Purification of the S1 protease 2 from *Saccharopolyspora erythraea***

The culture broth was centrifuged (20000 x g, 20 min) and the supernatant was carefully decanted from the precipitate. The supernatant was filtered through a Nalgene 0.2µm filtration unit in order to remove the rest of the *Bacillus* host cells. The 0.2µm filtrate was transferred to 20mM Tris/HCl, 1mM CaCl₂, pH 8.0 on a G25 sephadex column (from GE Healthcare). The G25
10 sephadex transferred filtrate was applied to a Q-sepharose FF column (from GE Healthcare) equilibrated in 20mM Tris/HCl, 1mM CaCl₂, pH 8.0. After washing the column with the equilibration buffer, the protease was eluted with a linear NaCl gradient (0 → 0.5M) in the same buffer over five column volumes. Fractions from the column were analysed for protease activity (using the Suc-AAPF-pNA assay at pH 9) and peak-fractions were pooled. Solid (NH₄)₂SO₄ was added to the pool
15 from the Q-sepharose FF column to a final 2.0M (NH₄)₂SO₄ concentration and the salt adjusted pool was applied to a Phenyl-Toyopearl column (from TosoHaas) equilibrated in 100mM H₃BO₃, 10mM MES, 2mM CaCl₂, 2.0M (NH₄)₂SO₄, pH 6.0. After washing the column with the equilibration buffer, the protease was eluted with a linear (NH₄)₂SO₄ gradient (2.0 → 0M) in the same buffer over five column volumes. Fractions from the column were analysed for protease activity (using the
20 Suc-AAPF-pNA assay at pH 9) and peak-fractions were pooled. The pool from the Phenyl-Toyopearl column was diluted 3 times with 50mM Citric acid/NaOH, 1mM CaCl₂, pH 4.0 and the pH was adjusted to pH 4.0 with 1M HCl. The adjusted pool was applied to an XpressLine ProA column (from UpFront chromatography) equilibrated in 50mM Citric acid/NaOH, 1mM CaCl₂, pH 4.0. After washing the column extensively with the equilibration buffer, the protease was step-eluted with
25 50mM Tris/HCl, pH 9.0. Fractions from the column were analysed for protease activity (using the Suc-AAPF-pNA assay at pH 9) and active fractions were analysed by SDS-PAGE. The fractions, where only one band was seen on the coomassie stained SDS-PAGE gel, were pooled and were used for further characterization.

30 **Example 3: Characterization of the S1 protease 2 from *Saccharopolyspora erythraea***

The Suc-AAPF-pNA assay was used for obtaining the pH-activity profile and the pH-stability profile (residual activity after 2 hours at indicated pH-values). For the pH-stability profile the protease was diluted 10x in the different assay buffers to reach the pH-values of these buffers and then incubated for 2 hours at 37°C. After incubation, the pH of the protease incubations was

adjusted to the same pH-value by dilution in the pH 9.0 assay buffer. Residual activities were measured at pH 9.0 relative to a sample, which was kept at stable conditions (5°C, pH 9.0). The Protazyme AK assay was used for obtaining the temperature-activity profile at pH 7.0. The Suc-AAPX-pNA assay and ten different Suc-AAPX-pNA substrates were used for obtaining the P1-specificity of the enzymes at pH 9.0. The results are shown in tables 2-5 and figures 1-4.

Table 2: pH-activity profile of at 37°C as determined using the Suc-AAPF-pNA assay

pH	S1 protease 2 from <i>Saccharopolyspora erythraea</i>	Protease 10R
2	0.00	-
3	0.00	0.00
4	0.01	0.02
5	0.06	0.07
6	0.25	0.21
7	0.56	0.44
8	0.86	0.67
9	1.00	0.88
10	0.95	1.00
11	0.84	0.93

Note: activities are relative to the optimal pH for the enzyme.

Table 3: pH-stability profile (residual activity after 2 hours at 37°C) as determined using the Suc-AAPF-pNA assay

pH	S1 protease 2 from <i>Saccharopolyspora erythraea</i>	Protease 10R
2	0.04	0.78
3	0.21	1.03
4	0.93	0.99
5	1.01	1.00
6	0.98	1.03
7	0.98	1.01
8	0.96	0.98
9	0.94	0.99
10	0.88	0.99
11	0.22	0.86
After 2 hours at 5 °C	1.00 (at pH 9)	1.00 (at pH 9)

Note: activities are residual activities relative to a sample, which was kept at stable conditions (5°C, pH 9.0).

Table 4: Temperature activity profile as determined using the Protazyme AK assay

Temp (°C)	S1 protease 2 from <i>Saccharopolyspora erythraea</i> (pH 7)	Protease 10R (pH 6.5)
15	0.12	0.01
25	0.18	0.02
37	0.45	0.06
50	0.81	0.13
60	1.00	0.35
70	0.33	0.96
80	-	1.00
90	-	0.18

Note: activities are relative to the optimal temperature at pH 7.0 or 6.5 for the enzyme.

Table 5: P1-specificity on 10 Suc-AAPX-pNA substrates at 25°C as determined using the Suc-AAPX-pNA assay

Suc-AAPX-pNA	S1 protease 2 from <i>Saccharopolyspora erythraea</i> (pH 9)	Protease 10R (pH 9)
Suc-AAPA-pNA	0.05	0.13
Suc-AAPR-pNA	0.11	0.09
Suc-AAPD-pNA	0.00	0.00
Suc-AAPI-pNA	0.00	0.00
Suc-AAPM-pNA	0.47	0.78
Suc-AAPV-pNA	0.00	0.01
Suc-AAPL-pNA	0.22	0.18
Suc-AAPE-pNA	0.00	0.00
Suc-AAPK-pNA	0.08	0.08
Suc-AAPF-pNA	1.00	1.00

Note: activities are relative to the best substrate (Suc-AAPF-pNA) for the enzyme.

5 Other characteristics for the S1 protease 2 from *Saccharopolyspora erythraea*

Inhibitors: PMSF and Cl-2A.

Determination of the N-terminal sequence by EDMAN degradation was: ADVIGGD.

The relative molecular weight as determined by SDS-PAGE was approx. $M_r = 21\text{kDa}$.

The molecular weight determined by intact molecular weight analysis was 18379.4Da.

10 The mature sequence (from MS data, EDMAN degradation data and DNA sequence):

ADVIGGDAYYIGSGSRCSVGFVQGGFVTAGHCGNQGDSTSQPSGTFEGSSFPGNDYGWVRTA
SGENPVPLVNDYQGGTVGVAGSSEAAEGASICRSGSTTGWHCGTVEAKNQTVRYPQGTVEGLT
RTNVCAEPGDSGGSWLSGDQAQGVTSGGSGDCTSGGTTYFQPVNEILQAYGLTLLTQ

The calculated molecular weight from this mature sequence was 18379.6Da.

15 **Example 4: Soybean-maize meal activity assay**

An end-point assay using soybean-maize meal as substrate was used for obtaining the pH activity profile of the proteases at pH 3-7.

Substrate: Soybean meal-maize meal mixed in a 30:70 ratio.

Assay buffers: 9 buffers containing 100 mM succinic acid, 100 mM HEPES, 100 mM CHES, 100 mM CAPS, 1 mM CaCl₂, 150 mM KCl, 0.01% Triton X-100 were prepared and adjusted using HCl or NaOH to a pH value such that after soybean-maize meal substrate (1 g) had been mixed with assay buffer (10 mL) to give a slurry, the final pH of the slurry was one of the following pH's: 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0 and 11.0.

Substrate slurry (2 mL) was mixed for 30 min before protease addition and incubation for 3 hours at 40°C (500 rpm). Protease (200 mg enzyme protein/kg dry matter) was dissolved in 100 µl 100 mM sodium acetate buffer (9.565 g/L NaOAc, 1.75 g/L acetic acid, 5 mM CaCl₂, 0.01% BSA, 0.01% Tween20, pH 6.0) and added. Samples were centrifuged (10 min, 4000 rpm, 0°C) and the supernatants collected for analysis using the o-Phthaldialdehyde (OPA) assay.

The results are shown in Table 6 below and figure 5. The proteolytic activity of the S1 protease 2 from *Saccharopolyspora erythraea* on soybean-maize meal increases with increasing pH from pH 3 to pH 7. While the S1 protease 2 from *Saccharopolyspora erythraea* is as active as Protease 10R at pH 6-7, it shows somewhat higher activity at pH 3-5 indicating that the S1 protease 2 from *Saccharopolyspora erythraea* might have the potential to be more efficient at hydrolyzing protein along the entire gastro-intestinal tract of monogastric animals.

Table 6: Protease activity (OD₃₄₀ x dilution factor) on soybean-maize meal at pH 3.0 to 7.0 at 40°C

pH	S1 protease 2 from <i>Saccharopolyspora erythraea</i>		Protease 10R	
	Average	Standard deviation	Average	Standard deviation
3.0	0.36	0.01	0.22	0.06
4.0	0.54	0.03	0.30	0.10
5.0	0.87	0.01	0.71	0.01
6.0	1.76	0.05	1.81	0.14
7.0	2.90	0.02	2.92	0.11

Figure 5 shows the activity (OD₃₄₀ x dilution factor) on soybean-maize meal of S1 protease 2 from *Saccharopolyspora erythraea* compared to the 10R protease.

Example 5: Thermostability

An aliquot of the protein sample of protease (purified as described in Example 2) is either desalted or buffer-changed into 20 mM Na-acetate, pH 4.0 using a prepacked PD-10 column or dialysed against 2 x 500 ml 20 mM Na-acetate, pH 4.0 at 4°C in a 2-3h step followed by an overnight step. The sample is 0.45 µm filtered and diluted with buffer to approx. 2 A280 units. The dialysis buffer is used as reference in Differential Scanning Calorimetry (DSC). The samples are degassed using vacuum suction and stirring for approx. 10 minutes.

A DSC scan is performed on a MicroCal VP-DSC at a constant scan rate of 1.5 °C/min from 20-90 °C. Data-handling is performed using the MicroCal Origin software (version 4.10), and the denaturation temperature, T_d (also called the melting temperature, T_m) is defined as the temperature at the apex of the peak in the thermogram.

Example 6: Steam stability

Residual activity of the protease after steam treatment may be evaluated using the following assay.

In these experiments a modified set-up is used whereby the steam is provided from a steam generator and led into the box. The samples placed on a plate are inserted into the box through a drawer when the temperature has reached ca. 93-94°C. Upon the insertion of the samples the temperature drops 4 °C. Incubation is performed for 30 seconds while the temperature remains approximately constant at 90°C. Thereafter the plate is quickly removed from the box, the samples placed on ice, re-suspended and evaluated with respect to protease activity using the Suc-AAPF-pNA or *o*-Phthaldialdehyde (OPA) assay. Each enzyme sample is compared to a similar sample that had not been steam treated in order to calculate residual activity.

Example 7: Pelleting stability tests

The enzyme granulation is performed in a manner as described in US No. patent 4,106,991, Example 1. The obtained granulate is dried in a fluid bed to a water content below 1% and sifted to obtain a product with the particle range 250 µm to 850 µm. Finally, the product is coated with palm oil and calcium carbonate in a manner as described in US patent No. 4,106,991, Example 22.

Approximately 50 g enzyme granulate is pre-mixed with 10 kg feed for 10 minutes in a small horizontal mixer. This premix is mixed with 90 kg feed for 10 minutes in a larger horizontal mixer. From the mixer the feed is led to the conditioner (a cascade mixer with steam injection) at a rate of approximately 300 kg/hour. The conditioner heats up the feed to 95°C (measured at the outlet) by injecting steam. The residence time in the conditioner is 30 seconds. From the conditioner the feed is led to a Simon Heesen press equipped with 3.0x35 mm horizontal die and pressed to pellets with

a length of around 15 mm. After the press the pellets are placed in an air cooler and cooled for 15 minutes.

The protease activity is measured using the Suc-AAPF-pNA assay prior to pelleting and in the feed pellets after pelleting. Pelleting stability is determined by comparing the protease activity
5 in pelleted feed relative to the activity in non-pelleted feed.

CLAIMS

What is claimed is:

1. Use of an isolated polypeptide having protease activity, selected from the group consisting of:
- 5 (a) a polypeptide having at least 80%, *e.g.*, 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%, at least 99%, or 100% sequence identity to SEQ ID NO: 5;
- 10 (b) a polypeptide encoded by a polynucleotide that hybridizes under medium-high stringency conditions, high stringency conditions, or very high stringency conditions with:
- (i) the mature polypeptide coding sequence of SEQ ID NO: 1;
- (ii) the mature polypeptide coding sequence of SEQ ID NO: 3; and/or
- (iii) the full-length complementary strand of (i) or (ii);
- 15 (c) a polypeptide encoded by a polynucleotide having at least 80%, *e.g.*, 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%, at least 99%, or 100% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 and/or SEQ ID NO: 3;
- 20 (d) a variant of the polypeptide of SEQ ID NO: 5 having at least 80% sequence identity to the polypeptide of SEQ ID NO: 5 comprising a substitution, deletion, and/or insertion at one or more (several) positions; and
- (e) a fragment of a polypeptide of (a), (b), (c) or (d), that has protease activity;
- in animal feed.
- 25
2. The use of claim 1, wherein the polypeptide comprises or consists of SEQ ID NO: 2.
3. The use of claim 1, wherein the polypeptide comprises or consists of SEQ ID NO: 4.
- 30 4. The use of any of claims 1-3, wherein the polypeptide is a fragment of SEQ ID NO: 2, SEQ ID NO: 4 or SEQ ID NO: 5, wherein the fragment has protease activity.

5. A variant polypeptide having protease activity and having at least 80%, *e.g.* 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 at least 99% sequence identity to the polypeptide of SEQ ID NO: 5, selected from the group consisting of:
- 5
- (a) a variant of the polypeptide of SEQ ID NO: 5 comprising at least one substitution, deletion, and/or insertion at one or more (several) positions; and
 - (b) a fragment of a polypeptide of (a) that has protease activity.
- 10 6. A composition comprising an isolated polypeptide having protease activity, selected from the group consisting of:
- (a) a polypeptide having at least 80%, *e.g.*, 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%, at least 99%, or 15 100% sequence identity to SEQ ID NO: 5;
 - (b) a polypeptide encoded by a polynucleotide that hybridizes under medium-high stringency conditions, high stringency conditions, or very high stringency conditions with:
 - (i) the mature polypeptide coding sequence of SEQ ID NO: 1;
 - 20 (ii) the mature polypeptide coding sequence of SEQ ID NO: 3; and/or
 - (iii) the full-length complementary strand of (i) or (ii);
 - (c) a polypeptide encoded by a polynucleotide having at least 80%, *e.g.*, 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%, at least 99%, or 100% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 and/or SEQ ID NO: 3;
 - 25 (d) a variant comprising a substitution, deletion, and/or insertion of one or more (several) amino acids of SEQ ID NO: 5; and
 - (e) a fragment of a polypeptide of (a), (b), (c) or (d), that has protease activity.
- 30
7. An isolated polynucleotide encoding the polypeptide indicated in any of claims 1-6 with the proviso that it is not 100% identical to SEQ ID NO: 1 or the mature polypeptide coding part thereof.

8. A nucleic acid construct or expression vector comprising the polynucleotide of claim 7 operably linked to one or more (several) control sequences that direct the production of the polypeptide in an expression host.
- 5 9. A recombinant host cell comprising the polynucleotide of claim 7 operably linked to one or more control sequences that direct the production of the polypeptide.
10. The host cell of claim 9, wherein the host is a bacterium, such as a *Bacillus*; a fungus, such as *Aspergillus*; or a yeast, such as *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*,
10 *Schizosaccharomyces*, or *Yarrowia*.
11. The host cell of claim 9, wherein the host is a *Bacillus*, such as *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus clausii*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, or *Bacillus thuringiensis*.
- 15 12. The host cell of claim 9, wherein the host is a *Streptomyces*, such as *Streptomyces lividans*, *Streptomyces coelicolor*, *Streptomyces avermitilis*, or *Streptomyces griseus*.
13. A method of producing the polypeptide of any of claims 1-6, comprising:
20 (a) cultivating a cell, which in its wild-type form produces the polypeptide, under conditions conducive for production of the polypeptide; and
(b) recovering the polypeptide.
14. A method of producing a polypeptide having protease activity, comprising:
25 (a) cultivating the host cell of any of claims 9-12 under conditions conducive for production of the polypeptide; and
(b) recovering the polypeptide.
15. A transgenic plant, plant part or plant cell transformed with a polynucleotide encoding the
30 polypeptide of any of claims 1-6.
16. A method of producing a polypeptide having protease activity, comprising:
(a) cultivating the transgenic plant or a plant cell of claim 15 under conditions conducive for production of the polypeptide; and
35 (b) recovering the polypeptide.

17. Use of at least one polypeptide indicated in any one of claims 1-6:
in animal feed;
in animal feed additives;
5 in the preparation of a composition for use in animal feed;
for improving the nutritional value of an animal feed;
for increasing digestible and/or soluble protein in animal feed;
for increasing the degree of hydrolysis of proteins in animal diets; and/or
for the treatment of proteins.
10
18. A method for improving the nutritional value of an animal feed, wherein at least one polypeptide indicated in any one of claims 1-6 is added to the feed.
19. An animal feed additive comprising:
15 at least one polypeptide indicated in any one of claims 1-6; and
at least one fat-soluble vitamin, and/or
at least one water-soluble vitamin, and/or
at least one trace mineral.
- 20 20. The animal feed additive of claim 19, which further comprises one or more amylases; phytases; xylanases; galactanases; alpha-galactosidases; proteases, phospholipases, beta-glucanases, or any mixture thereof.
21. An animal feed comprising an animal feed additive of claim 18 or 19.
25
22. The animal feed according to claim 21 having a crude protein content of 50 to 800 g/kg.
23. A method for the treatment of proteins, comprising the step of adding at least one polypeptide indicated in any one of claims 1-6 to at least one protein or protein source.
30
24. The method of claim 23, wherein soybean is included amongst the at least one protein source.
25. Use of at least one polypeptide of any one of claims 1-6 in animal feed.
35

26. An animal feed composition comprising at least one polypeptide indicated in any of claims 1 to 6.
27. The animal feed composition of claim 26, wherein the composition comprises one or more
5 further enzymes.
28. The animal feed composition of claim 26, wherein the further enzymes are selected from the group comprising proteases, amylases, lipases, cutinases, cellulases, endoglucanases, xyloglucanases, pectinases, pectin lyases, xanthanases, peroxidases, haloperoxygenases,
10 catalases and mannanases, or any mixture thereof.

1/3

pH-activity

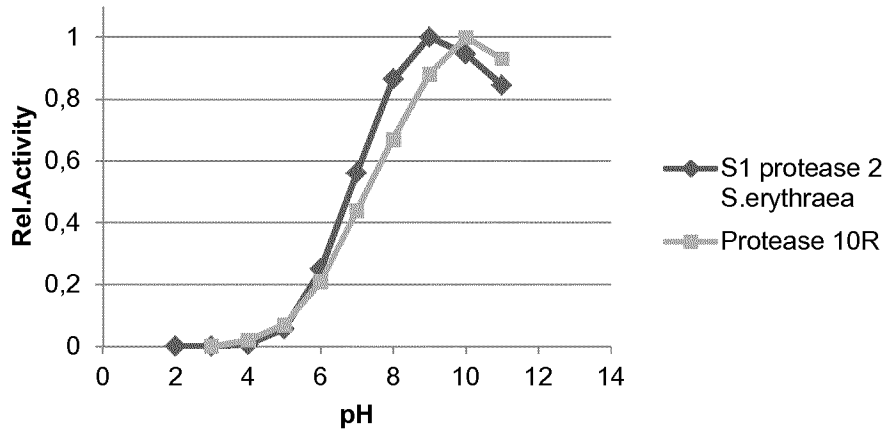


Figure 1

pH-stability. 2h at 37°C.

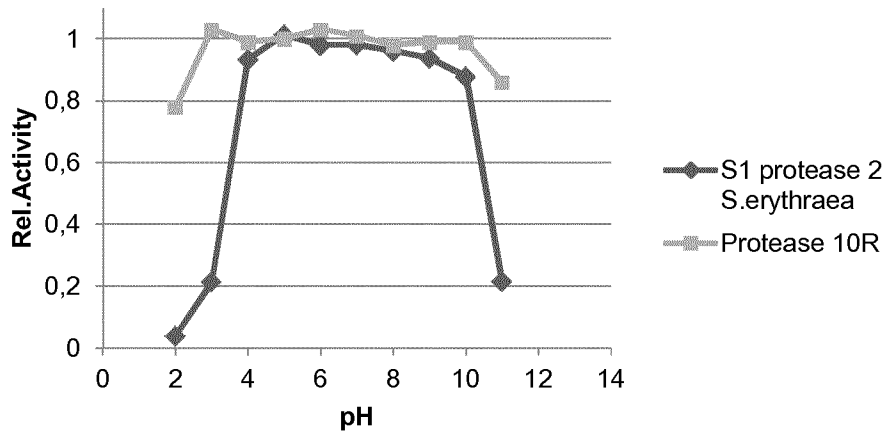


Figure 2

2/3

Temperature-activity

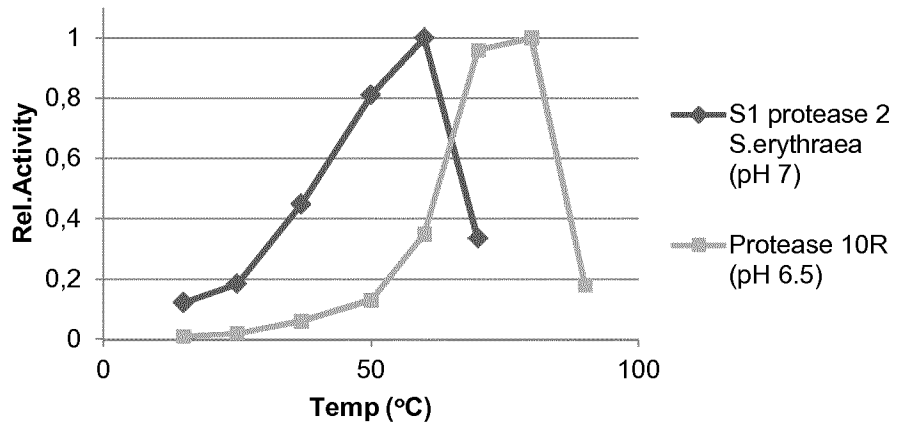


Figure 3

P1-specificity. Suc-AAPX-pNA.

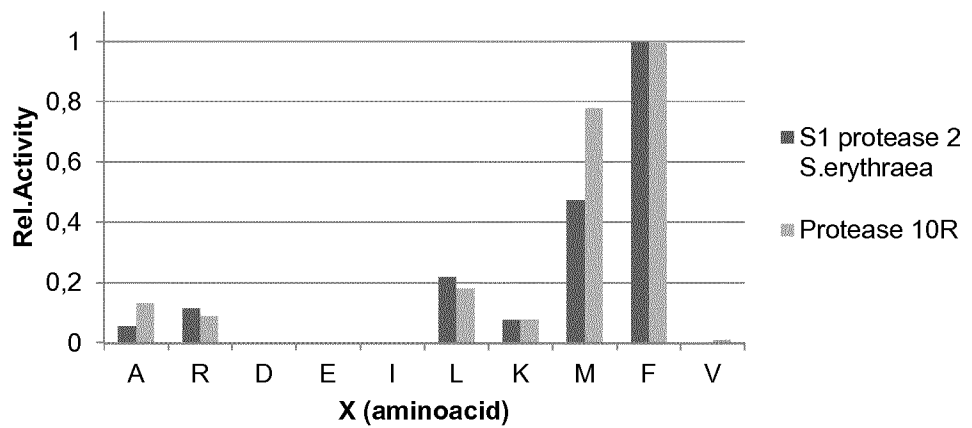


Figure 4

3/3

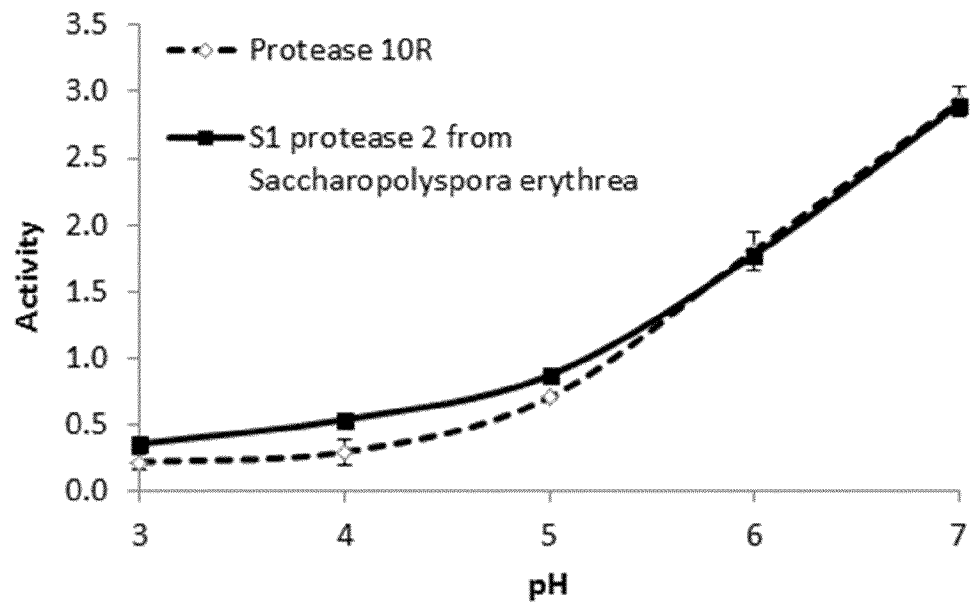


Figure 5