



- (51) International Patent Classification: *C12N 9/54* (2006.01) *C11D 3/386* (2006.01)
- (74) Agent: COHEN, Jacqueline; Danisco US Inc., 925 Page Mill Road., Palo Alto, US 94304 (US).
- (21) International Application Number: PCT/US2015/057512
- (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, 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.
- (22) International Filing Date: 27 October 2015 (27.10.2015)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 62/069,179 27 October 2014 (27.10.2014) US
- (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, ST, 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,
- (71) Applicant: DANISCO US INC. [US/US]; 925 Page Mill Road, Palo Alto, California 94304 (US).
- (72) Inventors: KOLKMAN, Marc; Linnaeushof 25, NL-2341 PA Oegstgeest (NL). KELLETT-SMITH, Anja Hemmingsen; Langelinieparken 3C, st. lejl 2, DK-8000 Århus C (DK). MEJLDAL, Rie; Ydingvej 11, DK-8752 Østbirk (DK). BABE, Lilia Maria; Emerald Hills, California (US).

[Continued on next page]

(54) Title: SERINE PROTEASES OF *BACILLUS SPECIES*

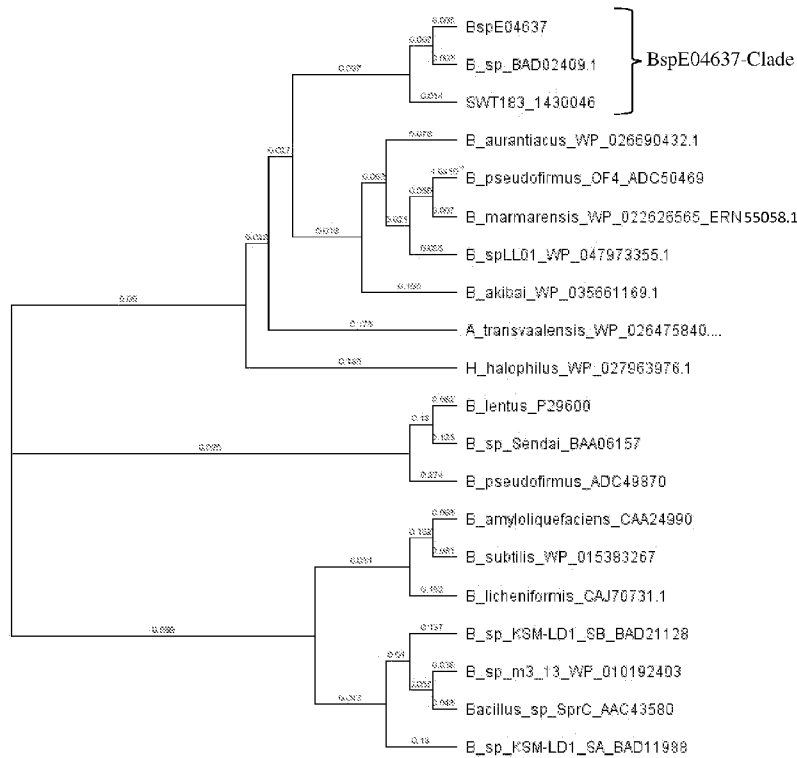


Figure 4

(57) Abstract: The present disclosure relates to serine proteases cloned from *Bacillus spp.*, and variants thereof. Compositions containing the serine proteases are suitable for use in cleaning fabrics and hard surfaces, as well as in a variety of industrial applications.

WO 2016/069557 A1

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

**Published:**

— *with international search report (Art. 21(3))*

## SERINE PROTEASES OF *BACILLUS SPECIES*

[001] The present disclosure relates to serine proteases cloned from *Bacillus spp.*, and variants thereof. Compositions containing the serine proteases are suitable for use in cleaning fabrics and hard surfaces, as well as in a variety of industrial applications.

[002] Serine proteases are enzymes (EC No. 3.4.21) possessing an active site serine that initiates hydrolysis of peptide bonds of proteins. There are two broad categories of serine proteases, based on their structure: chymotrypsin-like (trypsin-like) and subtilisin-like. The prototypical subtilisin (EC No. 3.4.21.62) was initially obtained from *Bacillus subtilis*. Subtilisins and their homologues are members of the S8 peptidase family of the MEROPS classification scheme. Members of family S8 have a catalytic triad in the order Asp, His and Ser in their amino acid sequence.

[003] Although serine proteases have long been known in the art of industrial enzymes, there remains a need for further serine proteases that are suitable for particular conditions and uses.

[004] The present compositions and methods relate to recombinant serine proteases cloned from *Bacillus spp.*, and variants thereof. Compositions containing the serine proteases are suitable for use in cleaning fabrics and hard surfaces, as well as in a variety of industrial applications.

[005] In some embodiments, the invention is a BspE04637-clade of subtilisins. In some embodiments, the invention is a recombinant polypeptide or active fragment thereof of a BspE04637-clade subtilisin. In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGI XXXHXDLXXXGGXSVFXXXXXXXXXXXXDXXGH (SEQ ID NO:31) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid ("Motif 1"). In other embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDL XXXGGXSVFXXXXXXXXDPXXDXXGH (SEQ ID NO:32) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid ("Motif 2"). In still further embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDL NVXGGXSVFXXXXXXXXXXXXDXXGH (SEQ ID NO:33) motif, wherein the initial D is the

active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid (“Motif 3”). In yet another embodiment, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDL NVXGGXSVFXXXXXXXXDPXXDXXGH (SEQ ID NO:34) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid (“Motif 4”). In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIDXXHXDLNVXGGXSVFXXXXXXXXXXDXXGH (SEQ ID NO:35) motif, wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid (“Motif 5”). In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIDXNHXDLNVRGGXSVFTXXXXX DPXXDXXGH (SEQ ID NO:36) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid (“Motif 6”). In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIDXNHXDLNVRGGXSVFTXXXXX DPYYDXXGH (SEQ ID NO:37) motif, wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid (“Motif 7”).

**[006]** In some embodiments, the invention is a recombinant polypeptide or an active fragment thereof comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99% amino acid sequence identity to an amino acid sequence of SEQ ID NO:3, 6 or 9. In some embodiments, the recombinant polypeptide or active fragment thereof comprises an amino acid sequence having at least 80% amino acid sequence identity to the amino acid sequence of SEQ ID NO:3, 6 or 9. In some embodiments, the recombinant polypeptide or active fragment thereof comprises an amino acid sequence having at least 80% amino acid sequence identity to the amino acid sequence of SEQ ID NO:3, 6 or 9, with the proviso that the polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001.

**[007]** In some embodiments, the recombinant polypeptide has protease activity, or subtilisin activity, specifically casein hydrolysis activity. In some embodiments, the recombinant polypeptide retains at least 50% of its maximal protease activity at a pH range of 6 to 12. In some embodiments, the recombinant polypeptide retains at least 50% of its maximal protease

activity at a temperature range of 60°C to 80°C. In some embodiments, the recombinant polypeptide has cleaning activity in a detergent composition, including an automatic dish washing detergent and a laundry detergent.

[008] In some embodiments, the invention is a composition comprising a surfactant and the recombinant polypeptide stated above. In some embodiments, the surfactant is selected from the group consisting of a non-ionic surfactant, an anionic surfactant, a cationic surfactant, a zwitterionic surfactant, an ampholytic surfactant, a semi-polar non-ionic surfactant, and a combination thereof. In some embodiments, the composition is a detergent composition, such as a laundry detergent, a fabric softening detergent, a dishwashing detergent, and a hard-surface cleaning detergent. In some embodiments, the composition further comprises at least one calcium ion and/or zinc ion, at least one stabilizer, at least one bleaching agent, phosphate, or borate. In some embodiments the composition is phosphate-free and/or borate-free. In some embodiments, the composition is a granular, powder, solid, bar, liquid, tablet, gel, paste or unit dose composition. In some embodiments, the composition further comprising one or more additional enzymes or enzyme derivatives selected from the group consisting of acyl transferases, alpha-amylases, beta-amylases, alpha-galactosidases, arabinosidases, aryl esterases, beta-galactosidases, carrageenases, catalases, cellobiohydrolases, cellulases, chondroitinases, cutinases, endo-beta-1, 4-glucanases, endo-beta-mannanases, esterases, exo-mannanases, galactanases, glucoamylases, hemicellulases, hyaluronidases, keratinases, laccases, lactases, ligninases, lipases, lipoxygenases, mannanases, oxidases, pectate lyases, pectin acetyl esterases, pectinases, pentosanases, peroxidases, phenoloxidases, phosphatases, phospholipases, phytases, polygalacturonases, proteases, pullulanases, reductases, rhamnogalacturonases, beta-glucanases, tannases, transglutaminases, xylan acetyl-esterases, xylanases, xyloglucanases, xylosidases, metalloproteases, additional serine proteases, and combinations thereof.

[009] In some embodiments, the invention is a method of cleaning, comprising contacting a surface or an item with a composition listed above. In some embodiments, the invention is a method for producing a recombinant polypeptide comprising stably transforming a host cell with an expression vector comprising a polynucleotide encoding the recombinant polypeptide above.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

[0010] Figure 1 provides a plasmid map for expression of BspE04637 protease.

[0011] Figure 2 provides a plot of protease activity of BspE04637 protease on DMC

substrate.

[0012] Figures 3A to 3F provide a MUSCLE multiple sequence alignment of subtilisins including BspE04637 and SWT183\_1430046.

[0013] Figure 4 provides a phylogenetic tree of subtilisins including BspE04637 and SWT183\_1430046 proteases.

[0014] Figure 5 illustrates the potential structural consequences of sequence motif changes found in the BspE04637-clade of subtilisins.

[0015] Described are compositions and methods relating to recombinant serine proteases from *Bacillus* species. The compositions and methods are based, in part, on the observation that recombinant BspE04637 and SWT\_1430046, among others, have protease activity in the presence of a surfactant, in basic reaction conditions, and at elevated temperatures. These features of BspE04637 and SWT\_1430046 make these proteases well suited for use in cleansing fabrics and hard surfaces, as well as in textile, leather and feather processing. The new proteases are also well suited to inclusion in compositions for protein degradation, including but not limited to laundry and dish washing detergents.

[0016] Prior to describing the present compositions and methods in detail, the following terms are defined for clarity. Terms and abbreviations not defined should be accorded their ordinary meaning as used in the art. Unless defined otherwise herein, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art. Unless otherwise indicated, the practice of the present disclosure involves conventional techniques commonly used in molecular biology, protein engineering, and microbiology. Although any methods and materials similar or equivalent to those described herein find use in the practice of the present disclosure, some suitable methods and materials are described herein. The terms defined immediately below are more fully described by reference to the Specification as a whole.

[0017] As used herein, the singular "a," "an" and "the" includes the plural unless the context clearly indicates otherwise. Unless otherwise indicated, nucleic acid sequences are written left to right in 5' to 3' orientation; and amino acid sequences are written left to right in amino to carboxy orientation. It is to be understood that this disclosure is not limited to the particular methodology, protocols, and reagents described herein, absent an indication to the contrary.

[0018] It is intended that every maximum numerical limitation given throughout this

Specification includes every lower numerical limitation, as if such lower numerical limitations were expressly written herein. Every minimum numerical limitation given throughout this Specification will include every higher numerical limitation, as if such higher numerical limitations were expressly written herein. Every numerical range given throughout this Specification will include every narrower numerical range that falls within such broader numerical range, as if such narrower numerical ranges were all expressly written herein.

**[0019]** As used herein in connection with a numerical value, the term “about” refers to a range of +/- 0.5 of the numerical value, unless the term is otherwise specifically defined in context. For instance, the phrase a “pH value of about 6” refers to pH values of from 5.5 to 6.5, unless the pH value is specifically defined otherwise.

**[0020]** As used herein, the terms “protease” and “proteinase” refer to an enzyme that has the ability to break down proteins and peptides. A protease has the ability to conduct “proteolysis,” by hydrolysis of peptide bonds that link amino acids together in a peptide or polypeptide chain forming the protein. This activity of a protease as a protein-digesting enzyme is referred to as “proteolytic activity.” Many well-known procedures exist for measuring proteolytic activity. For example, proteolytic activity may be ascertained by comparative assays that analyze the respective protease’s ability to hydrolyze a suitable substrate. Exemplary substrates useful in the analysis of protease or proteolytic activity, include, but are not limited to, di-methyl casein (Sigma C-9801), bovine collagen (Sigma C-9879), bovine elastin (Sigma E-1625), and bovine keratin (ICN Biomedical 902111). Colorimetric assays utilizing these substrates are well known in the art (See e.g., WO 99/34011 and U.S. Pat. No. 6,376,450). The pNA peptidyl assay (See e.g., Del Mar et al., *Anal Biochem*, 99:316-320, 1979) also finds use in determining the active enzyme concentration. This assay measures the rate at which p-nitroaniline is released as the enzyme hydrolyzes a soluble synthetic substrate, such as succinyl-alanine-alanine-proline-phenylalanine-p-nitroanilide (suc-AAPF-pNA). The rate of production of yellow color from the hydrolysis reaction is measured at 410 nm on a spectrophotometer and is proportional to the active enzyme concentration. In addition, absorbance measurements at 280 nanometers (nm) can be used to determine the total protein concentration in a sample of purified protein. The activity on substrate/protein concentration gives the enzyme specific activity.

**[0021]** The term “variant,” with respect to a polypeptide, refers to a polypeptide that differs from a specified wild-type, parental, or reference polypeptide in that it includes one or more

naturally-occurring or man-made substitutions, insertions, or deletions of an amino acid. Similarly, the term “variant,” with respect to a polynucleotide, refers to a polynucleotide that differs in nucleotide sequence from a specified wild-type, parental, or reference polynucleotide. The identity of the wild-type, parental, or reference polypeptide or polynucleotide will be apparent from context.

**[0022]** As used herein, “the genus *Bacillus*” includes all species within the genus “*Bacillus*,” as known to those of skill in the art, including but not limited to *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. gibsonii*, and *B. thuringiensis*. It is recognized that the genus *Bacillus* continues to undergo taxonomical reorganization. Thus, it is intended that the genus include species that have been reclassified, including but not limited to such organisms as *B. stearothermophilus*, which is now named “*Geobacillus stearothermophilus*”, or *B. polymyxa*, which is now “*Paenibacillus polymyxa*” The production of resistant endospores under stressful environmental conditions is considered the defining feature of the genus *Bacillus*, although this characteristic also applies to the recently named *Alicyclobacillus*, *Amphibacillus*, *Aneurinibacillus*, *Anoxybacillus*, *Brevibacillus*, *Filobacillus*, *Gracilibacillus*, *Halobacillus*, *Paenibacillus*, *Salibacillus*, *Thermobacillus*, *Ureibacillus*, and *Virgibacillus*.

**[0023]** As used herein, the term “mutation” refers to changes made to a reference amino acid or nucleic acid sequence. It is intended that the term encompass substitutions, insertions and deletions.

**[0024]** As used herein, the term “vector” refers to a nucleic acid construct used to introduce or transfer nucleic acid(s) into a target cell or tissue. A vector is typically used to introduce foreign DNA into a cell or tissue. Vectors include plasmids, cloning vectors, bacteriophages, viruses (e.g., viral vector), cosmids, expression vectors, shuttle vectors, and the like. A vector typically includes an origin of replication, a multicloning site, and a selectable marker. The process of inserting a vector into a target cell is typically referred to as transformation. The present invention includes, in some embodiments, a vector that comprises a DNA sequence encoding a serine protease polypeptide (e.g., precursor or mature serine protease polypeptide) that is operably linked to a suitable prosequence (e.g., secretory, signal peptide sequence, etc.) capable of effecting the expression of the DNA sequence in a suitable host, and the folding and



translocation of the recombinant polypeptide chain.

**[0025]** As used herein in the context of introducing a nucleic acid sequence into a cell, the term “introduced” refers to any method suitable for transferring the nucleic acid sequence into the cell. Such methods for introduction include but are not limited to protoplast fusion, transfection, transformation, electroporation, conjugation, and transduction. Transformation refers to the genetic alteration of a cell which results from the uptake, optional genomic incorporation, and expression of genetic material (e.g., DNA).

**[0026]** As used herein, a nucleic acid is “operably linked” with another nucleic acid sequence when it is placed into a functional relationship with another nucleic acid sequence. For example, a promoter or enhancer is operably linked to a nucleotide coding sequence if the promoter affects the transcription of the coding sequence. A ribosome binding site may be operably linked to a coding sequence if it is positioned so as to facilitate translation of the coding sequence. Typically, “operably linked” DNA sequences are contiguous. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers may be used in accordance with conventional practice.

**[0027]** As used herein the term “gene” refers to a polynucleotide (e.g., a DNA segment), that encodes a polypeptide and includes regions preceding and following the coding regions. In some instances a gene includes intervening sequences (introns) between individual coding segments (exons).

**[0028]** As used herein, “recombinant” when used with reference to a cell typically indicates that the cell has been modified by the introduction of a foreign nucleic acid sequence or that the cell is derived from a cell so modified. For example, a recombinant cell may comprise a gene not found in identical form within the native (non-recombinant) form of the cell, or a recombinant cell may comprise a native gene (found in the native form of the cell) that has been modified and re-introduced into the cell. A recombinant cell may comprise a nucleic acid endogenous to the cell that has been modified without removing the nucleic acid from the cell; such modifications include those obtained by gene replacement, site-specific mutation, and related techniques known to those of ordinary skill in the art. Recombinant DNA technology includes techniques for the production of recombinant DNA in vitro and transfer of the recombinant DNA into cells where it may be expressed or propagated, thereby producing a

recombinant polypeptide. “Recombination” and “recombining” of polynucleotides or nucleic acids refer generally to the assembly or combining of two or more nucleic acid or polynucleotide strands or fragments to generate a new polynucleotide or nucleic acid.

**[0029]** A nucleic acid or polynucleotide is said to “encode” a polypeptide if, in its native state or when manipulated by methods known to those of skill in the art, it can be transcribed and/or translated to produce the polypeptide or a fragment thereof. The anti-sense strand of such a nucleic acid is also said to encode the sequence.

**[0030]** The terms “host strain” and “host cell” refer to a suitable host for an expression vector comprising a DNA sequence of interest.

**[0031]** A “protein” or “polypeptide” comprises a polymeric sequence of amino acid residues. The terms “protein” and “polypeptide” are used interchangeably herein. The single and 3-letter code for amino acids as defined in conformity with the IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) is used throughout this disclosure. The single letter X refers to any of the twenty amino acids. It is also understood that a polypeptide may be coded for by more than one nucleotide sequence due to the degeneracy of the genetic code. Mutations can be named by the one letter code for the parent amino acid, followed by a position number and then the one letter code for the variant amino acid. For example, mutating glycine (G) at position 87 to serine (S) is represented as “G087S” or “G87S. When describing modifications, a position followed by amino acids listed in parentheses indicates a list of substitutions at that position by any of the listed amino acids. For example, 6(L,I) means position 6 can be substituted with a leucine or isoleucine. At times, in a sequence, a slash (/) is used to define substitutions, e.g. F/V, indicates that the particular position may have a phenylalanine or valine at that position.

**[0032]** A “prosequence” or “propeptide sequence” refers to an amino acid sequence between the signal peptide sequence and mature protease sequence that is necessary for the proper folding and secretion of the protease; they are sometimes referred to as intramolecular chaperones. Cleavage of the prosequence or propeptide sequence results in a mature active protease. Bacterial serine proteases are often expressed as pro-enzymes.

**[0033]** The terms “signal sequence” and “signal peptide” refer to a sequence of amino acid residues that may participate in the secretion or direct transport of the mature or precursor form of a protein. The signal sequence is typically located N-terminal to the precursor or mature protein sequence. The signal sequence may be endogenous or exogenous. A signal sequence is

normally absent from the mature protein. A signal sequence is typically cleaved from the protein by a signal peptidase after the protein is transported.

**[0034]** The term “mature” form of a protein, polypeptide, or peptide refers to the functional form of the protein, polypeptide, or peptide without the signal peptide sequence and propeptide sequence.

**[0035]** The term “precursor” form of a protein or peptide refers to a mature form of the protein having a prosequence operably linked to the amino or carbonyl terminus of the protein. The precursor may also have a “signal” sequence operably linked to the amino terminus of the prosequence. The precursor may also have additional polypeptides that are involved in post-translational activity (e.g., polypeptides cleaved therefrom to leave the mature form of a protein or peptide).

**[0036]** The term “wild-type” in reference to an amino acid sequence or nucleic acid sequence indicates that the amino acid sequence or nucleic acid sequence is a native or naturally-occurring sequence. As used herein, the term “naturally-occurring” refers to anything (e.g., proteins, amino acids, or nucleic acid sequences) that is found in nature. Conversely, the term “non-naturally occurring” refers to anything that is not found in nature (e.g., recombinant nucleic acids and protein sequences produced in the laboratory or modification of the wild-type sequence).

**[0037]** As used herein with regard to amino acid residue positions, “corresponding to” or “corresponds to” or “corresponds” refers to an amino acid residue at the enumerated position in a protein or peptide, or an amino acid residue that is analogous, homologous, or equivalent to an enumerated residue in a protein or peptide. As used herein, “corresponding region” generally refers to an analogous position in a related proteins or a reference protein.

**[0038]** The terms “derived from” and “obtained from” refer to not only a protein produced or producible by a strain of the organism in question, but also a protein encoded by a DNA sequence isolated from such strain and produced in a host organism containing such DNA sequence. Additionally, the term refers to a protein which is encoded by a DNA sequence of synthetic and/or cDNA origin and which has the identifying characteristics of the protein in question. To exemplify, “proteases derived from Bacillus” refers to those enzymes having proteolytic activity that are naturally produced by Bacillus, as well as to serine proteases like those produced by Bacillus sources but which through the use of genetic engineering techniques are produced by other host cells transformed with a nucleic acid encoding the serine proteases.

**[0039]** The term “identical” in the context of two polynucleotide or polypeptide sequences refers to the nucleic acids or amino acids in the two sequences that are the same when aligned for maximum correspondence, as measured using sequence comparison or analysis algorithms described below and known in the art.

**[0040]** As used herein, “% identity” or percent identity” or “PID” refers to protein sequence identity. Percent identity may be determined using standard techniques known in the art. Useful algorithms include the BLAST algorithms (See, Altschul et al., J Mol Biol, 215:403-410, 1990; and Karlin and Altschul, Proc Natl Acad Sci USA, 90:5873-5787, 1993). The BLAST program uses several search parameters, most of which are set to the default values. The NCBI BLAST algorithm finds the most relevant sequences in terms of biological similarity but is not recommended for query sequences of less than 20 residues (Altschul et al., Nucleic Acids Res, 25:3389-3402, 1997; and Schaffer et al., Nucleic Acids Res, 29:2994-3005, 2001). Exemplary default BLAST parameters for a nucleic acid sequence searches include: Neighboring words threshold=11; E-value cutoff=10; Scoring Matrix=NUC.3.1 (match=1, mismatch=-3); Gap Opening=5; and Gap Extension=2. Exemplary default BLAST parameters for amino acid sequence searches include: Word size = 3; E-value cutoff=10; Scoring Matrix=BLOSUM62; Gap Opening=11; and Gap extension=1. A percent (%) amino acid sequence identity value is determined by the number of matching identical residues divided by the total number of residues of the “reference” sequence including any gaps created by the program for optimal/maximum alignment. BLAST algorithms refer to the “reference” sequence as the “query” sequence.

**[0041]** As used herein, “homologous proteins” or “homologous proteases” refers to proteins that have distinct similarity in primary, secondary, and/or tertiary structure. Protein homology can refer to the similarity in linear amino acid sequence when proteins are aligned. Homologous search of protein sequences can be done using BLASTP and PSI-BLAST from NCBI BLAST with threshold (E-value cut-off) at 0.001. (Altschul SF, Madde TL, Shaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI BLAST a new generation of protein database search programs. Nucleic Acids Res 1997 Set 1;25(17):3389-402). Using this information, proteins sequences can be grouped. A phylogenetic tree can be built using the amino acid sequences. Amino acid sequences can be entered in a program such as the Vector NTI Advance suite and a Guide Tree can be created using the Neighbor Joining (NJ) method (Saitou and Nei, Mol Biol Evol, 4:406-425, 1987). The tree construction can be calculated using Kimura’s

correction for sequence distance and ignoring positions with gaps. A program such as AlignX can display the calculated distance values in parenthesis following the molecule name displayed on the phylogenetic tree.

**[0042]** Understanding the homology between molecules can reveal the evolutionary history of the molecules as well as information about their function; if a newly sequenced protein is homologous to an already characterized protein, there is a strong indication of the new protein's biochemical function. The most fundamental relationship between two entities is homology; two molecules are said to be homologous if they have been derived from a common ancestor. Homologous molecules, or homologs, can be divided into two classes, paralogs and orthologs. Paralogs are homologs that are present within one species. Paralogs often differ in their detailed biochemical functions. Orthologs are homologs that are present within different species and have very similar or identical functions. A protein superfamily is the largest grouping (clade) of proteins for which common ancestry can be inferred. Usually this common ancestry is based on sequence alignment and mechanistic similarity. Superfamilies typically contain several protein families which show sequence similarity within the family. The term “protein clan” is commonly used for protease superfamilies based on the MEROPS protease classification system.

**[0043]** The CLUSTAL W algorithm is another example of a sequence alignment algorithm (See, Thompson et al., *Nucleic Acids Res*, 22:4673-4680, 1994). Default parameters for the CLUSTAL W algorithm include: Gap opening penalty=10.0; Gap extension penalty=0.05; Protein weight matrix=BLOSUM series; DNA weight matrix=IUB; Delay divergent sequences %=40; Gap separation distance=8; DNA transitions weight=0.50; List hydrophilic residues=GPSNDQEKR; Use negative matrix=OFF; Toggle Residue specific penalties=ON; Toggle hydrophilic penalties=ON; and Toggle end gap separation penalty=OFF. In CLUSTAL algorithms, deletions occurring at either terminus are included. For example, a variant with a five amino acid deletion at either terminus (or within the polypeptide) of a polypeptide of 500 amino acids would have a percent sequence identity of 99% (495/500 identical residues × 100) relative to the “reference” polypeptide. Such a variant would be encompassed by a variant having “at least 99% sequence identity” to the polypeptide.

**[0044]** A nucleic acid or polynucleotide is “isolated” when it is at least partially or completely separated from other components, including but not limited to for example, other proteins, nucleic acids, cells, etc. Similarly, a polypeptide, protein or peptide is “isolated” when

it is at least partially or completely separated from other components, including but not limited to for example, other proteins, nucleic acids, cells, etc. On a molar basis, an isolated species is more abundant than are other species in a composition. For example, an isolated species may comprise at least about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or about 100% (on a molar basis) of all macromolecular species present. Preferably, the species of interest is purified to essential homogeneity (i.e., contaminant species cannot be detected in the composition by conventional detection methods). Purity and homogeneity can be determined using a number of techniques well known in the art, such as agarose or polyacrylamide gel electrophoresis of a nucleic acid or a protein sample, respectively, followed by visualization upon staining. If desired, a high-resolution technique, such as high performance liquid chromatography (HPLC) or a similar means can be utilized for purification of the material.

**[0045]** The term “purified” as applied to nucleic acids or polypeptides generally denotes a nucleic acid or polypeptide that is essentially free from other components as determined by analytical techniques well known in the art (e.g., a purified polypeptide or polynucleotide forms a discrete band in an electrophoretic gel, chromatographic eluate, and/or a media subjected to density gradient centrifugation). For example, a nucleic acid or polypeptide that gives rise to essentially one band in an electrophoretic gel is “purified.” A purified nucleic acid or polypeptide is at least about 50% pure, usually at least about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, about 99.5%, about 99.6%, about 99.7%, about 99.8% or more pure (e.g., percent by weight on a molar basis). In a related sense, a composition is enriched for a molecule when there is a substantial increase in the concentration of the molecule after application of a purification or enrichment technique. The term “enriched” refers to a compound, polypeptide, cell, nucleic acid, amino acid, or other specified material or component that is present in a composition at a relative or absolute concentration that is higher than a starting composition.

**[0046]** As used herein, the term “functional assay” refers to an assay that provides an indication of a protein’s activity. In some embodiments, the term refers to assay systems in which a protein is analyzed for its ability to function in its usual capacity. For example, in the

case of a protease, a functional assay involves determining the effectiveness of the protease to hydrolyze a proteinaceous substrate.

**[0047]** The term “cleaning activity” refers to a cleaning performance achieved by a serine protease polypeptide or reference protease under conditions prevailing during the proteolytic, hydrolyzing, cleaning, or other process of the disclosure. In some embodiments, cleaning performance of a serine protease polypeptide or reference protease may be determined by using various assays for cleaning one or more various enzyme sensitive stains on an item or surface (e.g., a stain resulting from food, grass, blood, ink, milk, oil, and/or egg protein). Cleaning performance of a variant or reference protease can be determined by subjecting the stain on the item or surface to standard wash condition(s) and assessing the degree to which the stain is removed by using various chromatographic, spectrophotometric, or other quantitative methodologies. Exemplary cleaning assays and methods are known in the art and include, but are not limited to those described in WO99/34011 and US 6,605,458, both of which are herein incorporated by reference, as well as those cleaning assays and methods included in the Examples provided below.

**[0048]** The term “cleaning effective amount” of a serine protease polypeptide or reference protease refers to the amount of protease that achieves a desired level of enzymatic activity in a specific cleaning composition. Such effective amounts are readily ascertained by one of ordinary skill in the art and are based on many factors, such as the particular protease used, the cleaning application, the specific composition of the cleaning composition, and whether a liquid or dry (e.g., granular, tablet, bar) composition is required, etc.

**[0049]** The term “cleaning adjunct material” refers to any liquid, solid, or gaseous material included in cleaning composition other than a serine protease polypeptide of the disclosure. In some embodiments, the cleaning compositions of the present disclosure include one or more cleaning adjunct materials. Each cleaning adjunct material is typically selected depending on the particular type and form of cleaning composition (e.g., liquid, granule, powder, bar, paste, spray, tablet, gel, foam, or other composition). Preferably, each cleaning adjunct material is compatible with the protease enzyme used in the composition.

**[0050]** Cleaning compositions and cleaning formulations include any composition that is suited for cleaning, bleaching, disinfecting, and/or sterilizing any object, item, and/or surface. Such compositions and formulations include, but are not limited to for example, liquid and/or

solid compositions, including cleaning or detergent compositions (e.g., liquid, tablet, gel, bar, granule, and/or solid laundry cleaning or detergent compositions and fine fabric detergent compositions; hard surface cleaning compositions and formulations, such as for glass, wood, ceramic and metal counter tops and windows; carpet cleaners; oven cleaners; fabric fresheners; fabric softeners; and textile, laundry booster cleaning or detergent compositions, laundry additive cleaning compositions, and laundry pre-spotter cleaning compositions; dishwashing compositions, including hand or manual dishwashing compositions (e.g., “hand” or “manual” dishwashing detergents) and automatic dishwashing compositions (e.g., “automatic dishwashing detergents”). Single dosage unit forms also find use with the present invention, including but not limited to pills, tablets, gencaps, or other single dosage units such as pre-measured powders or liquids.

**[0051]** Cleaning composition or cleaning formulations, as used herein, include, unless otherwise indicated, granular or powder-form all-purpose or heavy-duty washing agents, especially cleaning detergents; liquid, granular, gel, solid, tablet, paste, or unit dosage form all-purpose washing agents, especially the so-called heavy-duty liquid (HDL) detergent or heavy-duty dry (HDD) detergent types; liquid fine-fabric detergents; hand or manual dishwashing agents, including those of the high-foaming type; hand or manual dishwashing, automatic dishwashing, or dishware or tableware washing agents, including the various tablet, powder, solid, granular, liquid, gel, and rinse-aid types for household and institutional use; liquid cleaning and disinfecting agents, including antibacterial hand-wash types, cleaning bars, mouthwashes, denture cleaners, car shampoos, carpet shampoos, bathroom cleaners; hair shampoos and/or hair-rinses for humans and other animals; shower gels and foam baths and metal cleaners; as well as cleaning auxiliaries, such as bleach additives and “stain-stick” or pre-treat types. In some embodiments, granular compositions are in “compact” form; in some embodiments, liquid compositions are in a “concentrated” form.

**[0052]** As used herein, “fabric cleaning compositions” include hand and machine laundry detergent compositions including laundry additive compositions and compositions suitable for use in the soaking and/or pretreatment of stained fabrics (e.g., clothes, linens, and other textile materials).

**[0053]** As used herein, “non-fabric cleaning compositions” include non-textile (i.e., non-fabric) surface cleaning compositions, including, but not limited to for example, hand or manual



or automatic dishwashing detergent compositions, oral cleaning compositions, denture cleaning compositions, contact lens cleaning compositions, wound debridement compositions, and personal cleansing compositions.

**[0054]** As used herein, the term “detergent composition” or “detergent formulation” is used in reference to a composition intended for use in a wash medium for the cleaning of soiled or dirty objects, including particular fabric and/or non-fabric objects or items. Such compositions of the present disclosure are not limited to any particular detergent composition or formulation. Indeed, in some embodiments, the detergents of the disclosure comprise at least one serine protease polypeptide of the disclosure and, in addition, one or more surfactants, transferase(s), hydrolytic enzymes, oxido reductases, builders (e.g., a builder salt), bleaching agents, bleach activators, bluing agents, fluorescent dyes, caking inhibitors, masking agents, enzyme activators, antioxidants, and/or solubilizers. In some instances, a builder salt is a mixture of a silicate salt and a phosphate salt, preferably with more silicate (e.g., sodium metasilicate) than phosphate (e.g., sodium tripolyphosphate). Some compositions of the disclosure, such as, but not limited to, cleaning compositions or detergent compositions, do not contain any phosphate (e.g., phosphate salt or phosphate builder).

**[0055]** As used herein, the term “bleaching” refers to the treatment of a material (e.g., fabric, laundry, pulp, etc.) or surface for a sufficient length of time and/or under appropriate pH and/or temperature conditions to effect a brightening (i.e., whitening) and/or cleaning of the material. Examples of chemicals suitable for bleaching include, but are not limited to, for example,  $\text{ClO}_2$ ,  $\text{H}_2\text{O}_2$ , peracids,  $\text{NO}_2$ , etc.

**[0056]** As used herein, “wash performance” of a protease (e.g., a serine protease polypeptide of the disclosure) refers to the contribution of a serine protease polypeptide to washing that provides additional cleaning performance to the detergent as compared to the detergent without the addition of the serine protease polypeptide to the composition. Wash performance is compared under relevant washing conditions. In some test systems, other relevant factors, such as detergent composition, sud concentration, water hardness, washing mechanics, time, pH, and/or temperature, can be controlled in such a way that condition(s) typical for household application in a certain market segment (e.g., hand or manual dishwashing, automatic dishwashing, dishware cleaning, tableware cleaning, fabric cleaning, etc.) are imitated.

**[0057]** The term “relevant washing conditions” is used herein to indicate the conditions,

particularly washing temperature, time, washing mechanics, sud concentration, type of detergent and water hardness, actually used in households in a hand dishwashing, automatic dishwashing, or laundry detergent market segment.

**[0058]** As used herein, the term “disinfecting” refers to the removal of contaminants from the surfaces, as well as the inhibition or killing of microbes on the surfaces of items. It is not intended that the present disclosure be limited to any particular surface, item, or contaminant(s) or microbes to be removed.

**[0059]** The “compact” form of the cleaning compositions herein is best reflected by density and, in terms of composition, by the amount of inorganic filler salt. Inorganic filler salts are conventional ingredients of detergent compositions in powder form. In conventional detergent compositions, the filler salts are present in substantial amounts, typically about 17 to about 35% by weight of the total composition. In contrast, in compact compositions, the filler salt is present in amounts not exceeding about 15% of the total composition. In some embodiments, the filler salt is present in amounts that do not exceed about 10%, or more preferably, about 5%, by weight of the composition. In some embodiments, the inorganic filler salts are selected from the alkali and alkaline-earth-metal salts of sulfates and chlorides. In some embodiments, the filler salt is sodium sulfate.

**[0060]** The present disclosure provides novel serine protease enzymes. The serine protease polypeptides of the present disclosure include isolated, recombinant, substantially pure, or non-naturally occurring polypeptides. In some embodiments, the polypeptides are useful in cleaning applications and can be incorporated into cleaning compositions that are useful in methods of cleaning an item or a surface in need thereof.

**[0061]** The BspE04637-clade of subtilisins is characterized by a 2 amino acid residue insertion after position N56, wherein the amino acid position is numbered by correspondence with the amino acid sequence of subtilisin BPN'. This insertion occurs in the span of residues linking the catalytic aspartic acid (D32) and catalytic histidine (H66) that form part of the characteristic catalytic triad, wherein the amino acid positions are numbered by correspondence with the amino acid sequence of subtilisin BspE04637.

**[0062]** In some embodiments, the invention is a BspE04637-clade of subtilisins. In some embodiments, the invention is a recombinant polypeptide or active fragment thereof of a BspE04637-clade subtilisin. In some embodiments, the BspE04637-clade of subtilisins

comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDLXXXGGXSVFXXXXXXXXXXDXXGH (SEQ ID NO:31) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid (“Motif 1”). In other embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDLXXXGGXSVFXXXXXXXXDPXXDXXGH (SEQ ID NO:32) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid (“Motif 2”). In still further embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDLNVXGGXSVFXXXXXXXXXXDXXGH (SEQ ID NO:33) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid (“Motif 3”). In yet another embodiment, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDLNVXGGXSVFXXXXXXXXDPXXDXXGH (SEQ ID NO:34) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid (“Motif 4”). In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIDXXHXDLNVXGGXSVFXXXXXXXXXXDXXGH (SEQ ID NO:35) motif, wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid (“Motif 5”). In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIDXNHXDLNVRGGXSVFTXXXXXDPXXDXXGH (SEQ ID NO:36) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid (“Motif 6”). In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIDXNHXDLNVRGGXSVFTXXXXXDPYYDXXGH (SEQ ID NO:37) motif, wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid (“Motif 7”).

**[0063]** In some embodiments, the invention is a BspE04637-clade of subtilisins, with the proviso that the BspE04637-clade of subtilisins does not comprise BAD02409 or JP2003325186-0001 subtilisin. In some embodiments, the invention is a recombinant polypeptide or active fragment thereof of a BspE04637-clade subtilisin, with the proviso that the recombinant

polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001. In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDLXXXGGXSVFXXXXXX XXXXDXXGH (SEQ ID NO:31) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid, with the proviso that the subtilisin and/or recombinant polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001. In other embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDL XXXGGXSVFXXXXXXDPXXDXXGH (SEQ ID NO:32) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid, with the proviso that the subtilisin and/or recombinant polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001. In still further embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDLNVXGGXSVFXXXXXX XXXXDXXGH (SEQ ID NO:33) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid, with the proviso that the subtilisin and/or recombinant polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001. In yet another embodiment, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIXXXHXDLNVXGGXS VFXXXXXXDPXXDXXGH (SEQ ID NO:34) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid, with the proviso that the subtilisin and/or recombinant polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001. In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIDXXHXDLNVXGGXSVFXXXXXX XXXXDXXGH (SEQ ID NO:35) motif, wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid, with the proviso that the subtilisin and/or recombinant polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001. In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIDXNHXDLNVRGGXSVFTXXXXX DPXXDXXGH (SEQ ID NO:36) motif, wherein the

initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid, with the proviso that the subtilisin and/or recombinant polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001. In some embodiments, the BspE04637-clade of subtilisins comprises a subtilisin or recombinant polypeptide or active fragment thereof comprising a DTGIDXNHXDLNVRGGXSVFTXXXXDPYYDXXGH (SEQ ID NO:37) motif, wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid, with the proviso that the subtilisin and/or recombinant polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001.

**[0064]** In some embodiments, the polypeptide of the present invention is a polypeptide having a specified degree of amino acid sequence homology to the exemplified polypeptides, e.g., 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:3, 6 and 9. In some embodiments, the recombinant polypeptide or active fragment thereof comprises an amino acid sequence having at least 70% amino acid sequence identity to the amino acid sequence of SEQ ID NO:3, 6 or 9. In some embodiments, the recombinant polypeptide or active fragment thereof comprises an amino acid sequence having at least 70% amino acid sequence identity to the amino acid sequence of SEQ ID NO:3, 6 or 9, with the proviso that the polypeptide or active fragment thereof does not comprise BAD02409, WP\_026690432, ADC50469, ERN55058, WP\_022626565, WP\_026475840, WP\_027963976, WP\_012957833, or JP2003325186-0001. In some embodiments, the recombinant polypeptide or active fragment thereof comprises an amino acid sequence having at least 70% amino acid sequence identity to the amino acid sequence of SEQ ID NO:3, 6 or 9, with the proviso that the polypeptide or active fragment thereof does not comprise BAD02409, WP\_026690432, ADC50469, ERN55058, WP\_022626565, WP\_026475840, WP\_027963976, WP\_012957833, WP\_035661169, WP\_047973355, or JP2003325186-0001. In some embodiments, the recombinant polypeptide or active fragment thereof comprises an amino acid sequence having at least 75% amino acid sequence identity to the amino acid sequence of SEQ ID NO:3, 6 or 9. In some embodiments, the recombinant polypeptide or active fragment thereof comprises an amino acid sequence having at least 75% amino acid sequence identity to the amino acid sequence of SEQ ID NO:3, 6 or 9, with the proviso that the polypeptide or active fragment thereof does not comprise

BAD02409, WP\_026690432, ADC50469, ERN55058, WP\_022626565, or JP2003325186-0001. In some embodiments, the recombinant polypeptide or active fragment thereof comprises an amino acid sequence having at least 75% amino acid sequence identity to the amino acid sequence of SEQ ID NO:3, 6 or 9, with the proviso that the polypeptide or active fragment thereof does not comprise BAD02409, WP\_026690432, ADC50469, ERN55058, WP\_022626565, WP\_035661169, WP\_047973355, or JP2003325186-0001. In some embodiments, the recombinant polypeptide or active fragment thereof comprises an amino acid sequence having at least 80%, 90%, or 95% amino acid sequence identity to the amino acid sequence of SEQ ID NO:3, 6 or 9. In some embodiments, the recombinant polypeptide or active fragment thereof comprises an amino acid sequence having at least 80%, 90%, or 95% amino acid sequence identity to the amino acid sequence of SEQ ID NO:3, 6 or 9, with the proviso that the polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001. Homology can be determined by amino acid sequence alignment, e.g., using a program such as BLAST, ALIGN, or CLUSTAL, as described herein. In some embodiments, the polypeptide is an isolated, recombinant, substantially pure, or non-naturally occurring enzyme having protease activity, such as subtilisin activity, or casein hydrolysis activity (for example, dimethylcasein hydrolysis activity).

**[0065]** Also provided is a polypeptide enzyme of the present invention, having protease activity, such as alkaline protease activity, said enzyme comprising an amino acid sequence which differs from the amino acid sequence of SEQ ID NO:3, 6 or 9 by no more than 50, no more than 40, no more than 30, no more than 25, no more than 20, no more than 15, no more than 10, no more than 9, no more than 8, no more than 7, no more than 6, no more than 5, no more than 4, no more than 3, no more than 2, or no more than 1 amino acid residue(s), when aligned using any of the previously described alignment methods.

**[0066]** As noted above, the variant enzyme polypeptides of the invention have enzymatic activities (e.g., protease activities) and thus are useful in cleaning applications, including but not limited to, methods for cleaning dishware items, tableware items, fabrics, and items having hard surfaces (e.g., the hard surface of a table, table top, wall, furniture item, floor, ceiling, etc.). Exemplary cleaning compositions comprising one or more variant serine protease enzyme polypeptides of the invention are described infra. The enzymatic activity (e.g., protease enzyme activity) of an enzyme polypeptide of the invention can be determined readily using procedures

well known to those of ordinary skill in the art. The Examples presented infra describe methods for evaluating the enzymatic activity and cleaning performance. The performance of polypeptide enzymes of the invention in removing stains (e.g., a protein stain such as blood/milk/ink or egg yolk), cleaning hard surfaces, or cleaning laundry, dishware or tableware item(s) can be readily determined using procedures well known in the art and/or by using procedures set forth in the Examples.

**[0067]** The serine protease polypeptides of the present invention can have protease activity over a broad range of pH conditions. In some embodiments, the serine protease polypeptides have protease activity on dimethylcasein as a substrate, as demonstrated in Examples below. In some embodiments, the serine protease polypeptides have protease activity at a pH of from about 4.0 to about 12.0. In some embodiments, the serine protease polypeptides have protease activity at a pH of from about 6.0 to about 12.0. In some embodiments, the serine protease polypeptides have at least 50%, 60%, 70%, 80% or 90% of maximal protease activity at a pH of from about 6.0 to about 12.0, or from about 7.0 to about 12.0. In some embodiments, the serine protease polypeptides have protease activity at a pH above 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10.0, 10.5, 11.0 or 11.5. In some embodiments, the serine protease polypeptides have protease activity at a pH below 12.0, 11.5, 11.0, 10.5, 10.0, 9.5, 9.0, 8.5, 8.0, 7.5, 7.0, or 6.5.

**[0068]** In some embodiments, the serine protease polypeptides of the present invention have protease activity at a temperature range from about 10°C to about 90°C, or from about 30°C to about 80°C. In some embodiments, the serine protease polypeptides of the present invention have protease activity at a temperature range of from about 55°C to about 75°C. In some embodiments, the serine protease polypeptides have at least 50%, 60%, 70%, 80% or 90% of maximal protease activity at a temperature of from about 55°C to about 75°C. In some embodiments, the serine proteases have activity at a temperature above 50°C, 55°C, 60°C, 65°C, or 70°C. In some embodiments, the serine proteases have activity at a temperature below 75°C, 80°C, 70°C, 65°C, 60°C, or 55°C.

**[0069]** In some embodiments, the serine protease polypeptides of the present invention have at least 60% activity after 20 minutes at 50°C under stressed conditions. In some embodiments, the serine protease polypeptides of the present invention have at least 20% activity after 20 minutes at 60°C under stressed conditions. In some embodiments, the serine protease polypeptides of the present invention have at least 15% activity after 20 minutes at 75°C under

stressed conditions. The stressed conditions can be, for example, those shown in the Examples. In some embodiments, the stressed condition is in an LAS/EDTA assay, Tris/EDTA assay, or OMO HDL assay.

**[0070]** In some embodiments, the serine protease polypeptides of the present invention demonstrate cleaning performance in a cleaning composition. Cleaning compositions often include ingredients harmful to the stability and performance of enzymes, making cleaning compositions a harsh environment for enzymes, e.g. serine proteases, to retain function. Thus, it is not trivial for an enzyme to be put in a cleaning composition and expect enzymatic function (e.g. serine protease activity, such as demonstrated by cleaning performance). In some embodiments, the serine protease polypeptides of the present invention demonstrate cleaning performance in automatic dishwashing (ADW) detergent compositions. In some embodiments, the cleaning performance in ADW detergent compositions includes cleaning of egg yolk stains. In some embodiments, the serine protease polypeptides of the present invention demonstrate cleaning performance in laundry detergent compositions. In some embodiments, the cleaning performance in laundry detergent compositions includes cleaning of blood/milk/ink stains. In each of the cleaning compositions, the serine protease polypeptides of the present invention demonstrate cleaning performance with or without a bleach component.

**[0071]** A polypeptide of the invention can be subject to various changes, such as one or more amino acid insertions, deletions, and/or substitutions, either conservative or non-conservative, including where such changes do not substantially alter the enzymatic activity of the polypeptide. Similarly, a nucleic acid of the invention can also be subject to various changes, such as one or more substitutions of one or more nucleotides in one or more codons such that a particular codon encodes the same or a different amino acid, resulting in either a silent variation (e.g., when the encoded amino acid is not altered by the nucleotide mutation) or non-silent variation, one or more deletions of one or more nucleic acids (or codons) in the sequence, one or more additions or insertions of one or more nucleic acids (or codons) in the sequence, and/or cleavage of or one or more truncations of one or more nucleic acids (or codons) in the sequence. Many such changes in the nucleic acid sequence may not substantially alter the enzymatic activity of the resulting encoded polypeptide enzyme compared to the polypeptide enzyme encoded by the original nucleic acid sequence. A nucleic acid sequence of the invention can also be modified to include one or more codons that provide for optimum expression in an expression system (e.g.,



bacterial expression system), while, if desired, said one or more codons still encode the same amino acid(s).

**[0072]** The invention provides isolated, non-naturally occurring, or recombinant nucleic acids which may be collectively referred to as “nucleic acids of the invention” or “polynucleotides of the invention”, which encode polypeptides of the invention. Nucleic acids of the invention, including all described below, are useful in recombinant production (e.g., expression) of polypeptides of the invention, typically through expression of a plasmid expression vector comprising a sequence encoding the polypeptide of interest or fragment thereof. As discussed above, polypeptides include serine protease polypeptides having enzymatic activity (e.g., proteolytic activity) which are useful in cleaning applications and cleaning compositions for cleaning an item or a surface (e.g., surface of an item) in need of cleaning.

**[0073]** In some embodiments, the polynucleotide of the present invention is a polynucleotide having a specified degree of nucleic acid homology to the exemplified polynucleotide. In some embodiments, the polynucleotide has a nucleic acid sequence having at least 50, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% nucleic acid identity to SEQ ID NO:1, 4 or 7. In some embodiments, the polynucleotide comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1, 4, and 7. In some embodiments, the polynucleotide has a nucleic acid sequence that encodes a polypeptide or an active fragment thereof having at least 70, 75, 80, 85, 90, 95, 96, 97, 98, 99, or 100% amino acid sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs: 3, 6, and 9. In some embodiments, the polynucleotide has a nucleic acid sequence that encodes a polypeptide or an active fragment thereof having SEQ ID NO:3, 6, or 9. In other embodiments, the polynucleotide may also have a complementary nucleic acid sequence to a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1, 4, and 7. Homology can be determined by amino acid sequence alignment, e.g., using a program such as BLAST, ALIGN, or CLUSTAL, as described herein.

**[0074]** In some embodiments, the invention provides an isolated, recombinant, substantially pure, synthetically derived, or non-naturally occurring nucleic acid comprising a nucleotide sequence encoding any polypeptide (including any fusion protein, etc.) of the invention described above in the section entitled "Polypeptides of the Invention" and elsewhere herein. The invention also provides an isolated, recombinant, substantially pure, synthetically derived, or

non-naturally-occurring nucleic acid comprising a nucleotide sequence encoding a combination of two or more of any polypeptides of the invention described above and elsewhere herein. The present invention provides nucleic acids encoding a serine protease polypeptide of the present invention, wherein the serine protease polypeptide is a mature form having proteolytic activity. In some embodiments, the serine protease is expressed recombinantly with a homologous pro-peptide sequence. In other embodiments, the serine protease is expressed recombinantly with a heterologous pro-peptide sequence (e.g., GG36 pro-peptide sequence).

**[0075]** Nucleic acids of the invention can be generated by using any suitable synthesis, manipulation, and/or isolation techniques, or combinations thereof. For example, a polynucleotide of the invention may be produced using standard nucleic acid synthesis techniques, such as solid-phase synthesis techniques that are well-known to those skilled in the art. In such techniques, fragments of up to 50 or more nucleotide bases are typically synthesized, then joined (e.g., by enzymatic or chemical ligation methods) to form essentially any desired continuous nucleic acid sequence. The synthesis of the nucleic acids of the invention can be also facilitated by any suitable method known in the art, including but not limited to chemical synthesis using the classical phosphoramidite method (See e.g., Beaucage et al. *Tetrahedron Letters* 22:1859-69 [1981]); or the method described by Matthes et al. (See, Matthes et al., *EMBO J.* 3:801-805 [1984], as is typically practiced in automated synthetic methods. Nucleic acids of the invention also can be produced by using an automatic DNA synthesizer. Customized nucleic acids can be ordered from a variety of commercial sources (e.g., The Midland Certified Reagent Company, the Great American Gene Company, Operon Technologies Inc., and DNA2.0). Other techniques for synthesizing nucleic acids and related principles are known in the art (See e.g., Itakura et al., *Ann. Rev. Biochem.* 53:323 [1984]; and Itakura et al., *Science* 198:1056 [1984]).

**[0076]** As indicated above, recombinant DNA techniques useful in modification of nucleic acids are well known in the art. For example, techniques such as restriction endonuclease digestion, ligation, reverse transcription and cDNA production, and polymerase chain reaction (e.g., PCR) are known and readily employed by those of skill in the art. Nucleotides of the invention may also be obtained by screening cDNA libraries using one or more oligonucleotide probes that can hybridize to or PCR-amplify polynucleotides which encode a serine protease polypeptide polypeptide(s) of the invention. Procedures for screening and isolating cDNA

clones and PCR amplification procedures are well known to those of skill in the art and described in standard references known to those skilled in the art. Some nucleic acids of the invention can be obtained by altering a naturally occurring polynucleotide backbone (e.g., that encodes an enzyme or parent protease) by, for example, a known mutagenesis procedure (e.g., site-directed mutagenesis, site saturation mutagenesis, and in vitro recombination). A variety of methods are known in the art that are suitable for generating modified polynucleotides of the invention that encode serine protease polypeptides of the invention, including, but not limited to, for example, site-saturation mutagenesis, scanning mutagenesis, insertional mutagenesis, deletion mutagenesis, random mutagenesis, site-directed mutagenesis, and directed-evolution, as well as various other recombinatorial approaches.

**[0077]** The present invention provides vectors comprising at least one serine protease polynucleotide of the invention described herein (e.g., a polynucleotide encoding a serine protease polypeptide of the invention described herein), expression vectors or expression cassettes comprising at least one nucleic acid or polynucleotide of the invention, isolated, substantially pure, or recombinant DNA constructs comprising at least one nucleic acid or polynucleotide of the invention, isolated or recombinant cells comprising at least one polynucleotide of the invention, and compositions comprising one or more such vectors, nucleic acids, expression vectors, expression cassettes, DNA constructs, cells, cell cultures, or any combination or mixtures thereof.

**[0078]** In some embodiments, the invention provides recombinant cells comprising at least one vector (e.g., expression vector or DNA construct) of the invention which comprises at least one nucleic acid or polynucleotide of the invention. Some such recombinant cells are transformed or transfected with such at least one vector, although other methods are available and known in the art. Such cells are typically referred to as host cells. Some such cells comprise bacterial cells, including, but are not limited to *Bacillus* sp. cells, such as *B. subtilis* cells. The invention also provides recombinant cells (e.g., recombinant host cells) comprising at least one serine protease polypeptide of the invention.

**[0079]** In some embodiments, the invention provides a vector comprising a nucleic acid or polynucleotide of the invention. In some embodiments, the vector is an expression vector or expression cassette in which a polynucleotide sequence of the invention which encodes a serine protease polypeptide of the invention is operably linked to one or additional nucleic acid

segments required for efficient gene expression (e.g., a promoter operably linked to the polynucleotide of the invention which encodes a serine protease polypeptide of the invention). A vector may include a transcription terminator and/or a selection gene, such as an antibiotic resistance gene, that enables continuous cultural maintenance of plasmid-infected host cells by growth in antimicrobial-containing media.

**[0080]** An expression vector may be derived from plasmid or viral DNA, or in alternative embodiments, contains elements of both. Exemplary vectors include, but are not limited to pC194, pJH101, pE194, pHP13 (See, Harwood and Cutting [eds.], Chapter 3, Molecular Biological Methods for Bacillus, John Wiley & Sons [1990]; suitable replicating plasmids for *B. subtilis* include those listed on p. 92) See also, Perego, Integrational Vectors for Genetic Manipulations in *Bacillus subtilis*, in Sonenshein et al., [eds.] *Bacillus subtilis* and Other Gram-Positive Bacteria: Biochemistry, Physiology and Molecular Genetics, American Society for Microbiology, Washington, D.C. [1993], pp. 615-624), and p2JM103BBI.

**[0081]** For expression and production of a protein of interest (e.g., serine protease polypeptide) in a cell, at least one expression vector comprising at least one copy of a polynucleotide encoding the serine protease polypeptide, and in some instances comprising multiple copies, is transformed into the cell under conditions suitable for expression of the serine protease. In some embodiments of the present invention, a polynucleotide sequence encoding the serine protease polypeptide (as well as other sequences included in the vector) is integrated into the genome of the host cell, while in other embodiments, a plasmid vector comprising a polynucleotide sequence encoding the serine protease polypeptide remains as autonomous extra-chromosomal element within the cell. The invention provides both extrachromosomal nucleic acid elements as well as incoming nucleotide sequences that are integrated into the host cell genome. The vectors described herein are useful for production of the serine protease polypeptides of the invention. In some embodiments, a polynucleotide construct encoding the serine protease polypeptide is present on an integrating vector that enables the integration and optionally the amplification of the polynucleotide encoding the serine protease polypeptide into the host chromosome. Examples of sites for integration are well known to those skilled in the art. In some embodiments, transcription of a polynucleotide encoding a serine protease polypeptide of the invention is effectuated by a promoter that is the wild-type promoter for the selected precursor protease. In some other embodiments, the promoter is heterologous to the

precursor protease, but is functional in the host cell. Specifically, examples of suitable promoters for use in bacterial host cells include, but are not limited to, for example, the amyE, amyQ, amyL, pstS, sacB, pSPAC, pAprE, pVeg, pHpaII promoters, the promoter of the *B. stearothermophilus* maltogenic amylase gene, the *B. amyloliquefaciens* (BAN) amylase gene, the *B. subtilis* alkaline protease gene, the *B. clausii* alkaline protease gene the *B. pumilis* xylosidase gene, the *B. thuringiensis* cryIII<sub>A</sub>, and the *B. licheniformis* alpha-amylase gene. Additional promoters include, but are not limited to the A4 promoter, as well as phage Lambda PR or PL promoters, and the *E. coli* lac, trp or tac promoters.

**[0082]** Serine protease polypeptides of the present invention can be produced in host cells of any suitable microorganism, including bacteria and fungi. In some embodiments, serine protease polypeptides of the present invention can be produced in Gram-positive bacteria. In some embodiments, the host cells are *Bacillus spp.*, *Streptomyces spp.*, *Escherichia spp.*, *Aspergillus spp.*, *Trichoderma spp.*, *Pseudomonas spp.*, *Corynebacterium spp.*, *Saccharomyces spp.*, or *Pichia spp.* In some embodiments, the serine protease polypeptides are produced by *Bacillus sp.* host cells. Examples of *Bacillus sp.* host cells that find use in the production of the serine protease polypeptides of the invention include, but are not limited to *B. licheniformis*, *B. lentus*, *B. subtilis*, *B. amyloliquefaciens*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. coagulans*, *B. circulans*, *B. pumilis*, *B. thuringiensis*, *B. clausii*, and *B. megaterium*, as well as other organisms within the genus *Bacillus*. In some embodiments, *B. subtilis* host cells are used for production of serine protease polypeptides. US 5,264,366 and 4,760,025 (RE 34,606) describe various *Bacillus* host strains that can be used for producing serine protease polypeptide of the invention, although other suitable strains can be used.

**[0083]** Several bacterial strains that can be used to produce serine protease polypeptides of the invention include non-recombinant (i.e., wild-type) *Bacillus sp.* strains, as well as variants of naturally-occurring strains and/or recombinant strains. In some embodiments, the host strain is a recombinant strain, wherein a polynucleotide encoding a polypeptide of interest has been introduced into the host. In some embodiments, the host strain is a *B. subtilis* host strain and particularly a recombinant *B. subtilis* host strain. Numerous *B. subtilis* strains are known, including, but not limited to for example, 1A6 (ATCC 39085), 168 (1A01), SB19, W23, Ts85, B637, PB1753 through PB1758, PB3360, JH642, 1A243 (ATCC 39,087), ATCC 21332, ATCC 6051, MI113, DE100 (ATCC 39,094), GX4931, PBT 110, and PEP 211 strain (See e.g., Hoch et

al., *Genetics* 73:215–228 [1973]; See also, U.S. Patent Nos. 4,450,235 and 4,302,544, and EP 0134048, each of which is incorporated by reference in its entirety). The use of *B. subtilis* as an expression host cells is well known in the art (See e.g., Palva et al., *Gene* 19:81-87 [1982]; Fahnestock and Fischer, *J. Bacteriol.*, 165:796–804 [1986]; and Wang et al., *Gene* 69:39–47 [1988]).

**[0084]** In some embodiments, the *Bacillus* host cell is a *Bacillus sp.* that includes a mutation or deletion in at least one of the following genes, degU, degS, degR and degQ. In some embodiments, the mutation is in a degU gene, and in some embodiments the mutation is degU(Hy)32 (See e.g., Msadek et al., *J. Bacteriol.* 172:824-834 [1990]; and Olmos et al., *Mol. Gen. Genet.* 253:562–567 [1997]). In some embodiments, the *Bacillus* host comprises a mutation or deletion in scoC4 (See e.g., Caldwell et al., *J. Bacteriol.* 183:7329-7340 [2001]); spoIIE (See e.g., Arigoni et al., *Mol. Microbiol.* 31:1407-1415 [1999]); and/or oppA or other genes of the opp operon (See e.g., Perego et al., *Mol. Microbiol.* 5:173-185 [1991]). Indeed, it is contemplated that any mutation in the opp operon that causes the same phenotype as a mutation in the oppA gene will find use in some embodiments of the altered *Bacillus* strain of the invention. In some embodiments, these mutations occur alone, while in other embodiments, combinations of mutations are present. In some embodiments, an altered *Bacillus* host cell strain that can be used to produce a serine protease polypeptide of the invention is a *Bacillus* host strain that already includes a mutation in one or more of the above-mentioned genes. In addition, *Bacillus sp.* host cells that comprise mutation(s) and/or deletions of endogenous protease genes find use. In some embodiments, the *Bacillus* host cell comprises a deletion of the aprE and the nprE genes. In other embodiments, the *Bacillus sp.* host cell comprises a deletion of 5 protease genes, while in other embodiments, the *Bacillus sp.* host cell comprises a deletion of 9 protease genes (See e.g., US 2005/0202535, incorporated herein by reference).

**[0085]** Host cells are transformed with at least one nucleic acid encoding at least one serine protease polypeptide of the invention using any suitable method known in the art. Methods for introducing a nucleic acid (e.g., DNA) into *Bacillus* cells or *E. coli* cells utilizing plasmid DNA constructs or vectors and transforming such plasmid DNA constructs or vectors into such cells are well known. In some embodiments, the plasmids are subsequently isolated from *E. coli* cells and transformed into *Bacillus* cells. However, it is not essential to use intervening microorganisms such as *E. coli*, and in some embodiments, a DNA construct or vector is directly

introduced into a *Bacillus* host.

**[0086]** Those of skill in the art are well aware of suitable methods for introducing nucleic acid sequences of the invention into *Bacillus* cells (See e.g., Ferrari et al., "Genetics," in Harwood et al. [eds.], *Bacillus*, Plenum Publishing Corp. [1989], pp. 57-72; Saunders et al., *J. Bacteriol.* 157:718-726 [1984]; Hoch et al., *J. Bacteriol.* 93:1925-1937 [1967]; Mann et al., *Current Microbiol.* 13:131-135 [1986]; Holubova, *Folia Microbiol.* 30:97 [1985]; Chang et al., *Mol. Gen. Genet.* 168:11-115 [1979]; Vorobjeva et al., *FEMS Microbiol. Lett.* 7:261-263 [1980]; Smith et al., *Appl. Env. Microbiol.* 51:634 [1986]; Fisher et al., *Arch. Microbiol.* 139:213-217 [1981]; and McDonald, *J. Gen. Microbiol.* 130:203 [1984]). Indeed, such methods as transformation, including protoplast transformation and transfection, transduction, and protoplast fusion are well known and suited for use in the present invention. Methods known in the art to transform *Bacillus* cells include such methods as plasmid marker rescue transformation, which involves the uptake of a donor plasmid by competent cells carrying a partially homologous resident plasmid (See, Contente et al., *Plasmid* 2:555-571 [1979]; Haima et al., *Mol. Gen. Genet.* 223:185-191 [1990]; Weinrauch et al., *J. Bacteriol.* 154:1077-1087 [1983]; and Weinrauch et al., *J. Bacteriol.* 169:1205-1211 [1987]). In this method, the incoming donor plasmid recombines with the homologous region of the resident "helper" plasmid in a process that mimics chromosomal transformation.

**[0087]** In addition to commonly used methods, in some embodiments, host cells are directly transformed with a DNA construct or vector comprising a nucleic acid encoding a serine protease polypeptide of the invention (i.e., an intermediate cell is not used to amplify, or otherwise process, the DNA construct or vector prior to introduction into the host cell). Introduction of the DNA construct or vector of the invention into the host cell includes those physical and chemical methods known in the art to introduce a nucleic acid sequence (e.g., DNA sequence) into a host cell without insertion into the host genome. Such methods include, but are not limited to calcium chloride precipitation, electroporation, naked DNA, liposomes and the like. In additional embodiments, DNA constructs or vector are co-transformed with a plasmid, without being inserted into the plasmid. In further embodiments, a selective marker is deleted from the altered *Bacillus* strain by methods known in the art (See, Stahl et al., *J. Bacteriol.* 158:411-418 [1984]; and Palmeros et al., *Gene* 247:255-264 [2000]).

**[0088]** In some embodiments, the transformed cells of the present invention are cultured in

conventional nutrient media. The suitable specific culture conditions, such as temperature, pH and the like are known to those skilled in the art and are well described in the scientific literature. In some embodiments, the invention provides a culture (e.g., cell culture) comprising at least one serine protease polypeptide or at least one nucleic acid of the invention.

**[0089]** In some embodiments, host cells transformed with at least one polynucleotide sequence encoding at least one serine protease polypeptide of the invention are cultured in a suitable nutrient medium under conditions permitting the expression of the present protease, after which the resulting protease is recovered from the culture. In some embodiments, the protease produced by the cells is recovered from the culture medium by conventional procedures, including, but not limited to for example, separating the host cells from the medium by centrifugation or filtration, precipitating the proteinaceous components of the supernatant or filtrate by means of a salt (e.g., ammonium sulfate), chromatographic purification (e.g., ion exchange, gel filtration, affinity, etc.).

**[0090]** In some embodiments, a serine protease polypeptide produced by a recombinant host cell is secreted into the culture medium. A nucleic acid sequence that encodes a purification facilitating domain may be used to facilitate purification of proteins. A vector or DNA construct comprising a polynucleotide sequence encoding a serine protease polypeptide may further comprise a nucleic acid sequence encoding a purification facilitating domain to facilitate purification of the serine protease polypeptide (See e.g., Kroll et al., DNA Cell Biol. 12:441-53 [1993]). Such purification facilitating domains include, but are not limited to, for example, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals (See, Porath, Protein Expr. Purif. 3:263-281 [1992]), protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system. The inclusion of a cleavable linker sequence such as Factor XA or enterokinase (e.g., sequences available from Invitrogen, San Diego, CA) between the purification domain and the heterologous protein also find use to facilitate purification.

**[0091]** Assays for detecting and measuring the enzymatic activity of an enzyme, such as a serine protease polypeptide of the invention, are well known. Various assays for detecting and measuring activity of proteases (e.g., serine protease polypeptides of the invention), are also known to those of ordinary skill in the art. In particular, assays are available for measuring protease activity that are based on the release of acid-soluble peptides from casein or



hemoglobin, measured as absorbance at 280 nm or colorimetrically using the Folin method. Other exemplary assays involve the solubilization of chromogenic substrates (See e.g., Ward, "Proteinases," in Fogarty (ed.), *Microbial Enzymes and Biotechnology*, Applied Science, London, [1983], pp. 251-317). Other exemplary assays include, but are not limited to succinyl-Ala-Ala-Pro-Phe-para nitroanilide assay (suc-AAPF-pNA) and the 2,4,6-trinitrobenzene sulfonate sodium salt assay (TNBS assay). Numerous additional references known to those in the art provide suitable methods (See e.g., Wells et al., *Nucleic Acids Res.* 11:7911-7925 [1983]; Christianson et al., *Anal. Biochem.* 223:119 -129 [1994]; and Hsia et al., *Anal Biochem.* 242:221-227 [1999]).

**[0092]** A variety of methods can be used to determine the level of production of a mature protease (e.g., mature serine protease polypeptides of the present invention) in a host cell. Such methods include, but are not limited to, for example, methods that utilize either polyclonal or monoclonal antibodies specific for the protease. Exemplary methods include, but are not limited to enzyme-linked immunosorbent assays (ELISA), radioimmunoassays (RIA), fluorescent immunoassays (FIA), and fluorescent activated cell sorting (FACS). These and other assays are well known in the art (See e.g., Maddox et al., *J. Exp. Med.* 158:1211 [1983]).

**[0093]** In some other embodiments, the invention provides methods for making or producing a mature serine protease polypeptide of the invention. A mature serine protease polypeptide does not include a signal peptide or a propeptide sequence. Some methods comprise making or producing a serine protease polypeptide of the invention in a recombinant bacterial host cell, such as for example, a *Bacillus sp.* cell (e.g., a *B. subtilis* cell). In some embodiments, the invention provides a method of producing a serine protease polypeptide of the invention, the method comprising cultivating a recombinant host cell comprising a recombinant expression vector comprising a nucleic acid encoding a serine protease polypeptide of the invention under conditions conducive to the production of the serine protease polypeptide. Some such methods further comprise recovering the serine protease polypeptide from the culture.

**[0094]** In some embodiments the invention provides methods of producing a serine protease polypeptide of the invention, the methods comprising: (a) introducing a recombinant expression vector comprising a nucleic acid encoding a serine protease polypeptide of the invention into a population of cells (e.g., bacterial cells, such as *B. subtilis* cells); and (b) culturing the cells in a culture medium under conditions conducive to produce the serine protease polypeptide encoded

by the expression vector. Some such methods further comprise: (c) isolating the serine protease polypeptide from the cells or from the culture medium.

**[0095]** Unless otherwise noted, all component or composition levels provided herein are made in reference to the active level of that component or composition, and are exclusive of impurities, for example, residual solvents or by-products, which may be present in commercially available sources. Enzyme components weights are based on total active protein. All percentages and ratios are calculated by weight unless otherwise indicated. All percentages and ratios are calculated based on the total composition unless otherwise indicated. Compositions of the invention include cleaning compositions, such as detergent compositions. In the exemplified detergent compositions, the enzymes levels are expressed by pure enzyme by weight of the total composition and unless otherwise specified, the detergent ingredients are expressed by weight of the total compositions.

**[0096]** While not essential for the purposes of the present invention, the non-limiting list of adjuncts illustrated hereinafter are suitable for use in the instant cleaning compositions. In some embodiments, these adjuncts are incorporated for example, to assist or enhance cleaning performance, for treatment of the substrate to be cleaned, or to modify the aesthetics of the cleaning composition as is the case with perfumes, colorants, dyes or the like. It is understood that such adjuncts are in addition to the serine protease polypeptides of the present invention. The precise nature of these additional components, and levels of incorporation thereof, will depend on the physical form of the composition and the nature of the cleaning operation for which it is to be used. Suitable adjunct materials include, but are not limited to, bleach catalysts, other enzymes, enzyme stabilizing systems, chelants, optical brighteners, soil release polymers, dye transfer agents, dispersants, suds suppressors, dyes, perfumes, colorants, filler salts, , photoactivators, fluorescers, fabric conditioners, hydrolyzable surfactants, preservatives, anti-oxidants, anti-shrinkage agents, anti-wrinkle agents, germicides, fungicides, color speckles, silvercare, anti-tarnish and/or anti-corrosion agents, alkalinity sources, solubilizing agents, carriers, processing aids, pigments, and pH control agents, surfactants, builders, chelating agents, dye transfer inhibiting agents, deposition aids, dispersants, additional enzymes, and enzyme stabilizers, catalytic materials, bleach activators, bleach boosters, hydrogen peroxide, sources of hydrogen peroxide, preformed peracids, polymeric dispersing agents, clay soil removal/anti-redeposition agents, brighteners, suds suppressors, dyes, perfumes, structure elasticizing agents,

fabric softeners, carriers, hydrotropes, processing aids and/or pigments. In addition to the disclosure below, suitable examples of such other adjuncts and levels of use are found in US 5,576,282; 6,306,812; 6,326,348; 6,610,642; 6,605,458; 5,705,464; 5,710,115; 5,698,504; 5,695,679; 5,686,014 and 5,646,101 all of which are incorporated herein by reference. In embodiments in which the cleaning adjunct materials are not compatible with the serine protease polypeptides of the present invention in the cleaning compositions, then suitable methods of keeping the cleaning adjunct materials and the protease(s) separated (i.e., not in contact with each other) until combination of the two components is appropriate are used. Such separation methods include any suitable method known in the art (e.g., gelcaps, encapsulation, tablets, physical separation, etc.). The aforementioned adjunct ingredients may constitute the balance of the cleaning compositions of the present invention.

**[0097]** The cleaning compositions of the present invention are advantageously employed for example, in laundry applications, hard surface cleaning applications, dishwashing applications, including automatic dishwashing and hand dishwashing, as well as cosmetic applications such as dentures, teeth, hair and skin cleaning. The enzymes of the present invention are also suited for use in contact lens cleaning and wound debridement applications. In addition, due to the unique advantages of increased effectiveness in lower temperature solutions, the enzymes of the present invention are ideally suited for laundry applications. Furthermore, the enzymes of the present invention find use in granular and liquid compositions.

**[0098]** The serine protease polypeptides of the present invention also find use in cleaning additive products. In some embodiments, low temperature solution cleaning applications find use. In some embodiments, the present invention provides cleaning additive products including at least one enzyme of the present invention is ideally suited for inclusion in a wash process when additional bleaching effectiveness is desired. Such instances include, but are not limited to low temperature solution cleaning applications. In some embodiments, the additive product is in its simplest form, one or more proteases. In some embodiments, the additive is packaged in dosage form for addition to a cleaning process. In some embodiments, the additive is packaged in dosage form for addition to a cleaning process where a source of peroxygen is employed and increased bleaching effectiveness is desired. Any suitable single dosage unit form finds use with the present invention, including but not limited to pills, tablets, gelcaps, or other single dosage units such as pre-measured powders or liquids. In some embodiments, filler(s) or carrier

material(s) are included to increase the volume of such compositions. Suitable filler or carrier materials include, but are not limited to, various salts of sulfate, carbonate and silicate as well as talc, clay and the like. Suitable filler or carrier materials for liquid compositions include, but are not limited to water or low molecular weight primary and secondary alcohols including polyols and diols. Examples of such alcohols include, but are not limited to, methanol, ethanol, propanol and isopropanol. In some embodiments, the compositions contain from about 5% to about 90% of such materials. Acidic fillers find use to reduce pH. Alternatively, in some embodiments, the cleaning additive includes adjunct ingredients, as more fully described below.

**[0099]** The present cleaning compositions and cleaning additives require an effective amount of at least one of the serine protease polypeptides provided herein, alone or in combination with other proteases and/or additional enzymes. The required level of enzyme is achieved by the addition of one or more serine protease polypeptides of the present invention. Typically the present cleaning compositions comprise at least about 0.0001 weight percent, from about 0.0001 to about 10, from about 0.001 to about 1, or from about 0.01 to about 0.1 weight percent of at least one of the serine protease polypeptides of the present invention.

**[00100]** The cleaning compositions herein are typically formulated such that, during use in aqueous cleaning operations, the wash water will have a pH of from about 4.0 to about 11.5, or even from about 5.0 to about 11.5, or even from about 5.0 to about 8.0, or even from about 7.5 to about 10.5. Liquid product formulations are typically formulated to have a pH from about 3.0 to about 9.0 or even from about 3 to about 5. Granular laundry products are typically formulated to have a pH from about 9 to about 11. In some embodiments, the cleaning compositions of the present invention can be formulated to have an alkaline pH under wash conditions, such as a pH of from about 8.0 to about 12.0, or from about 8.5 to about 11.0, or from about 9.0 to about 11.0. In some embodiments, the cleaning compositions of the present invention can be formulated to have a neutral pH under wash conditions, such as a pH of from about 5.0 to about 8.0, or from about 5.5 to about 8.0, or from about 6.0 to about 8.0, or from about 6.0 to about 7.5. In some embodiments, the neutral pH conditions can be measured when the cleaning composition is dissolved 1:100 (wt:wt) in de-ionised water at 20°C, measured using a conventional pH meter. Techniques for controlling pH at recommended usage levels include the use of buffers, alkalis, acids, etc., and are well known to those skilled in the art.

**[00101]** In some embodiments, when the serine protease polypeptide (s) is/are employed in a

granular composition or liquid, it is desirable for the serine protease polypeptide to be in the form of an encapsulated particle to protect the serine protease polypeptide from other components of the granular composition during storage. In addition, encapsulation is also a means of controlling the availability of the serine protease polypeptide during the cleaning process. In some embodiments, encapsulation enhances the performance of the serine protease polypeptide (s) and/or additional enzymes. In this regard, the serine protease polypeptides of the present invention are encapsulated with any suitable encapsulating material known in the art. In some embodiments, the encapsulating material typically encapsulates at least part of the serine protease polypeptide (s) of the present invention. Typically, the encapsulating material is water-soluble and/or water-dispersible. In some embodiments, the encapsulating material has a glass transition temperature (T<sub>g</sub>) of 0°C or higher. Glass transition temperature is described in more detail in WO97/11151. The encapsulating material is typically selected from consisting of carbohydrates, natural or synthetic gums, chitin, chitosan, cellulose and cellulose derivatives, silicates, phosphates, borates, polyvinyl alcohol, polyethylene glycol, paraffin waxes, and combinations thereof. When the encapsulating material is a carbohydrate, it is typically selected from monosaccharides, oligosaccharides, polysaccharides, and combinations thereof. In some typical embodiments, the encapsulating material is a starch (See e.g., EP0922499; US 4,977,252; 5,354,559, and 5,935,826). In some embodiments, the encapsulating material is a microsphere made from plastic such as thermoplastics, acrylonitrile, methacrylonitrile, polyacrylonitrile, polymethacrylonitrile and mixtures thereof; commercially available microspheres that find use include, but are not limited to those supplied by EXPANCEL<sup>®</sup> (Stockviksverken, Sweden), PM 6545, PM 6550, PM 7220, PM 7228, EXTENDOSPHERES<sup>®</sup>, LUXSIL<sup>®</sup>, Q-CEL<sup>®</sup>, and SPHERICEL<sup>®</sup> (PQ Corp., Valley Forge, PA).

**[00102]** There are a variety of wash conditions including varying detergent formulations, wash water volumes, wash water temperatures, and lengths of wash time, to which proteases involved in washing are exposed. A low detergent concentration system includes detergents where less than about 800 ppm of the detergent components are present in the wash water. A medium detergent concentration includes detergents where between about 800 ppm and about 2000ppm of the detergent components are present in the wash water. A high detergent concentration system includes detergents where greater than about 2000 ppm of the detergent components are present in the wash water. In some embodiments, the “cold water washing” of

the present invention utilizes “cold water detergent” suitable for washing at temperatures from about 10°C to about 40°C, or from about 20°C to about 30°C, or from about 15°C to about 25°C, as well as all other combinations within the range of about 15°C to about 35°C, and all ranges within 10°C to 40°C.

**[00103]** Different geographies typically have different water hardness. Water hardness is usually described in terms of the grains per gallon mixed  $\text{Ca}^{2+}/\text{Mg}^{2+}$ . Hardness is a measure of the amount of calcium ( $\text{Ca}^{2+}$ ) and magnesium ( $\text{Mg}^{2+}$ ) in the water. Most water in the United States is hard, but the degree of hardness varies. Moderately hard (60-120 ppm) to hard (121-181 ppm) water has 60 to 181 parts per million.

Table I. Water Hardness

Water	Grains per gallon	Parts per million
Soft	less than 1.0	less than 17
Slightly hard	1.0 to 3.5	17 to 60
Moderately hard	3.5 to 7.0	60 to 120
Hard	7.0 to 10.5	120 to 180
Very hard	greater than 10.5	greater than 180

**[00104]** Accordingly, in some embodiments, the present invention provides serine protease polypeptides that show surprising wash performance in at least one set of wash conditions (e.g., water temperature, water hardness, and/or detergent concentration). In some embodiments, the serine protease polypeptides of the present invention are comparable in wash performance to other serine protease polypeptide proteases. In some embodiments of the present invention, the serine protease polypeptides provided herein exhibit enhanced oxidative stability, enhanced thermal stability, enhanced cleaning capabilities under various conditions, and/or enhanced chelator stability. In addition, the serine protease polypeptides of the present invention find use in cleaning compositions that do not include detergents, again either alone or in combination with builders and stabilizers.

**[00105]** In some embodiments of the present invention, the cleaning compositions comprise at least one serine protease polypeptide of the present invention at a level from about 0.00001 % to about 10% by weight of the composition and the balance (e.g., about 99.999% to about 90.0%) comprising cleaning adjunct materials by weight of composition. In some other embodiments of the present invention, the cleaning compositions of the present invention comprises at least one serine protease polypeptide at a level of about 0.0001 % to about 10%, about 0.001% to about

5%, about 0.001% to about 2%, about 0.005% to about 0.5% by weight of the composition and the balance of the cleaning composition (e.g., about 99.9999% to about 90.0%, about 99.999 % to about 98%, about 99.995% to about 99.5% by weight) comprising cleaning adjunct materials.

**[00106]** In some embodiments, the cleaning compositions of the present invention comprise one or more additional detergent enzymes, which provide cleaning performance and/or fabric care and/or dishwashing benefits. Examples of suitable enzymes include, but are not limited to, acyl transferases, alpha-amylases, beta-amylases, alpha-galactosidases, arabinosidases, aryl esterases, beta-galactosidases, carrageenases, catalases, cellobiohydrolases, cellulases, chondroitinases, cutinases, endo-beta-1, 4-glucanases, endo-beta-mannanases, esterases, exo-mannanases, galactanases, glucoamylases, hemicellulases, hyaluronidases, keratinases, laccases, lactases, ligninases, lipases, lipoxygenases, mannanases, oxidases, pectate lyases, pectin acetyl esterases, pectinases, pentosanases, peroxidases, phenoloxidases, phosphatases, phospholipases, phytases, polygalacturonases, proteases, pullulanases, reductases, rhamnogalacturonases, beta-glucanases, tannases, transglutaminases, xylan acetyl-esterases, xylanases, xyloglucanases, and xylosidases, or any combinations or mixtures thereof. In some embodiments, a combination of enzymes is used (i.e., a “cocktail”) comprising conventional applicable enzymes like protease, lipase, cutinase and/or cellulase in conjunction with amylase is used.

**[00107]** In addition to the serine protease polypeptides provided herein, any other suitable protease finds use in the compositions of the present invention. Suitable proteases include those of animal, vegetable or microbial origin. In some embodiments, microbial proteases are used. In some embodiments, chemically or genetically modified mutants are included. In some embodiments, the protease is a serine protease, preferably an alkaline microbial protease or a trypsin-like protease. Examples of alkaline proteases include subtilisins, especially those derived from *Bacillus* (e.g., subtilisin, *lentus*, *amyloliquefaciens*, subtilisin Carlsberg, subtilisin 309, subtilisin 147 and subtilisin 168). Additional examples include those mutant proteases described in US RE 34,606; 5,955,340; 5,700,676; 6,312,936; and 6,482,628, all of which are incorporated herein by reference. Additional protease examples include, but are not limited to trypsin (e.g., of porcine or bovine origin), and the *Fusarium* protease described in WO 89/06270. In some embodiments, commercially available protease enzymes that find use in the present invention include, but are not limited to MAXATASE<sup>®</sup>, MAXACAL<sup>™</sup>, MAXAPEM<sup>™</sup>, OPTICLEAN<sup>®</sup>, OPTIMASE<sup>®</sup>, PROPERASE<sup>®</sup>, PURAFECT<sup>®</sup>, PURAFECT<sup>®</sup> OXP, PURAMAX<sup>™</sup>,

EXCELLASE™, PREFERENZ™ proteases (e.g. P100, P110, P280), EFFECTENZ™ proteases (e.g. P1000, P1050, P2000), EXCELLENZ™ proteases (e.g. P1000), ULTIMASE®, and PURAFAST™ (Genencor); ALCALASE®, SAVINASE®, PRIMASE®, DURAZYM™, POLARZYME®, OVOZYME®, KANNASE®, LIQUANASE®, NEUTRASE®, RELEASE® and ESPERASE® (Novozymes); BLAP™ and BLAP™ variants (Henkel Kommanditgesellschaft auf Aktien, Duesseldorf, Germany), and KAP (*B. alkalophilus* subtilisin; Kao Corp., Tokyo, Japan). Various proteases are described in WO95/23221, WO92/21760, WO 09/149200, WO09/149144, WO09/149145, WO11/072099, WO10/056640, WO10/056653, WO 11/140364, WO12/151534, US 2008/0090747, and US 5,801,039; 5,340,735; 5,500,364; 5,855,625; RE 34,606; 5,955,340; 5,700,676; 6,312,936; 6,482,628; 8,530,219; and various other patents. In some further embodiments, neutral metalloproteases find use in the present invention, including but not limited to the neutral metalloproteases described in WO1999014341, WO1999033960, WO1999014342, WO 1999034003, WO2007044993, WO2009058303, WO2009058661, WO2014/071410, WO2014/194032, WO2014/194034, WO2014/194054, and WO2014/194117. Exemplary metalloproteases include nprE, the recombinant form of neutral metalloprotease expressed in *B. subtilis* (See e.g., WO 07/044993), and PMN, the purified neutral metalloprotease from *B. amyloliquefaciens*.

**[00108]** In addition, any suitable lipase finds use in the present invention. Suitable lipases include, but are not limited to those of bacterial or fungal origin. Chemically or genetically modified mutants are encompassed by the present invention. Examples of useful lipases include *H. lanuginosa* lipase (See e.g., EP258068, and EP305216), *Rhizomucor miehei* lipase (See e.g., EP238023), *Candida* lipase, such as *C. antarctica* lipase (e.g., the *C. antarctica* lipase A or B; See e.g., EP214761), *Pseudomonas* lipases such as *P. alcaligenes* lipase and *P. pseudoalcaligenes* lipase (See e.g., EP 218 272), *P. cepacia* lipase (See e.g., EP331376), *P. stutzeri* lipase (See e.g., GB 1,372,034), *P. fluorescens* lipase, *Bacillus* lipase (e.g., *B. subtilis* lipase [Dartois et al., Biochem. Biophys. Acta 1131:253-260 [1993]]; *B. stearothermophilus* lipase [See e.g., JP 64/744992]; and *B. pumilus* lipase [See e.g., WO91/16422]).

**[00109]** Furthermore, a number of cloned lipases find use in some embodiments of the present invention, including but not limited to *Penicillium camembertii* lipase (See, Yamaguchi et al., Gene 103:61-67 [1991]), *Geotricum candidum* lipase (See, Shimada et al., J. Biochem., 106:383-388 [1989]), and various *Rhizopus* lipases such as *R. delemar* lipase (See, Hass et al.,



Gene 109:117-113 [1991]), a *R. niveus* lipase (Kugimiya et al., Biosci. Biotech. Biochem. 56:716-719 [1992]) and *R. oryzae* lipase.

**[00110]** Other types of lipase polypeptide enzymes such as cutinases also find use in some embodiments of the present invention, including but not limited to the cutinase derived from *Pseudomonas mendocina* (See, WO88/09367), and the cutinase derived from *Fusarium solani pisi* (See, WO 90/09446).

**[00111]** Additional suitable lipases include lipases such as M1 LIPASE™, LUMA FAST™, and LIPOMAX™ (Genencor); LIPEX®, LIPOLASE® and LIPOLASE® ULTRA (Novozymes); and LIPASE P™ "Amano" (Amano Pharmaceutical Co. Ltd., Japan).

**[00112]** In some embodiments of the present invention, the cleaning compositions of the present invention further comprise lipases at a level from about 0.00001 % to about 10% of additional lipase by weight of the composition and the balance of cleaning adjunct materials by weight of composition. In some other embodiments of the present invention, the cleaning compositions of the present invention also comprise lipases at a level of about 0.0001 % to about 10%, about 0.001% to about 5%, about 0.001% to about 2%, about 0.005% to about 0.5% lipase by weight of the composition.

**[00113]** In some embodiments of the present invention, any suitable amylase finds use in the present invention. In some embodiments, any amylase (e.g., alpha and/or beta) suitable for use in alkaline solutions also find use. Suitable amylases include, but are not limited to those of bacterial or fungal origin. Chemically or genetically modified mutants are included in some embodiments. Amylases that find use in the present invention, include, but are not limited to  $\alpha$ -amylases obtained from *B. licheniformis* (See e.g., GB 1,296,839). Additional suitable amylases include those found in W09510603, W09526397, W09623874, W09623873, W09741213, W09919467, W00060060, W00029560, W09923211, W09946399, W00060058, W00060059, W09942567, W00114532, W002092797, W00166712, W00188107, W00196537, W00210355, W09402597, W00231124, W09943793, W09943794, W02004113551, W02005001064, W02005003311, W00164852, W02006063594, W02006066594, W02006066596, W02006012899, W02008092919, W02008000825, W02005018336, W02005066338, W02009140504, W02005019443, W02010091221, W02010088447, W00134784, W02006012902, W02006031554, W02006136161, W02008101894, W02010059413, W02011098531, W02011080352, W02011080353, W02011080354, W0

2011082425, WO2011082429, WO2011076123, WO2011087836, WO2011076897, WO 94183314, WO9535382, WO9909183, WO9826078, WO9902702, WO9743424, WO9929876, WO9100353, WO9605295, WO9630481, WO9710342, WO2008088493, WO2009149419, WO 2009061381, WO2009100102, WO2010104675, WO2010117511, and WO2010115021.

Commercially available amylases that find use in the present invention include, but are not limited to DURAMYL<sup>®</sup>, TERMAMYL<sup>®</sup>, FUNGAMYL<sup>®</sup>, STAINZYME<sup>®</sup>, STAINZYME PLUS<sup>®</sup>, STAINZYME ULTRA<sup>®</sup>, and BAN<sup>™</sup> (Novozymes), as well as POWERASE<sup>™</sup>, RAPIDASE<sup>®</sup> and MAXAMYL<sup>®</sup> P (Genencor).

**[00114]** In some embodiments of the present invention, the cleaning compositions of the present invention further comprise amylases at a level from about 0.00001 to about 10% of additional amylase by weight of the composition and the balance of cleaning adjunct materials by weight of composition. In some other embodiments of the present invention, the cleaning compositions of the present invention also comprise amylases at a level of about 0.0001 to about 10%, about 0.001 to about 5%, about 0.001 to about 2%, about 0.005 to about 0.5% amylase by weight of the composition.

**[00115]** In some further embodiments, any suitable cellulase finds used in the cleaning compositions of the present invention. Suitable cellulases include, but are not limited to those of bacterial or fungal origin. Chemically or genetically modified mutants are included in some embodiments. Suitable cellulases include, but are not limited to Humicola insolens cellulases (See e.g., US 4,435,307). Especially suitable cellulases are the cellulases having color care benefits (See e.g., EP0495257). Commercially available cellulases that find use in the present include, but are not limited to CELLUZYME<sup>®</sup>, CAREZYME<sup>®</sup> (Novozymes), REVITALENZ<sup>™</sup> 100 (Danisco US Inc) and KAC-500(B)<sup>™</sup> (Kao Corporation). In some embodiments, cellulases are incorporated as portions or fragments of mature wild-type or variant cellulases, wherein a portion of the N-terminus is deleted (See e.g., US 5,874,276). Additional suitable cellulases include those found in WO2005054475, WO2005056787, US 7,449,318 and 7,833,773. In some embodiments, the cleaning compositions of the present invention further comprise cellulases at a level from about 0.00001 to about 10% of additional cellulase by weight of the composition and the balance of cleaning adjunct materials by weight of composition. In some other embodiments of the present invention, the cleaning compositions of the present invention also comprise cellulases at a level of about 0.0001 to about 10%, about 0.001 to about 5%, about

0.001 to about 2%, about 0.005 to about 0.5% cellulase by weight of the composition.

**[00116]** Any mannanase suitable for use in detergent compositions also finds use in the present invention. Suitable mannanases include, but are not limited to those of bacterial or fungal origin. Chemically or genetically modified mutants are included in some embodiments. Various mannanases are known which find use in the present invention (See e.g., US 6,566,114; 6,602,842; and 6,440,991, all of which are incorporated herein by reference). Commercially available mannanases that find use in the present invention include, but are not limited to MANNASTAR<sup>®</sup>, PURABRITE<sup>™</sup>, and MANNAWAY<sup>®</sup>. In some embodiments, the cleaning compositions of the present invention further comprise mannanases at a level from about 0.00001 to about 10% of additional mannanase by weight of the composition and the balance of cleaning adjunct materials by weight of composition. In some embodiments of the present invention, the cleaning compositions of the present invention also comprise mannanases at a level of about 0.0001 to about 10%, about 0.001 to about 5%, about 0.001 to about 2%, about 0.005 to about 0.5% mannanase by weight of the composition.

**[00117]** In some embodiments, peroxidases are used in combination with hydrogen peroxide or a source thereof (e.g., a percarbonate, perborate or persulfate) in the compositions of the present invention. In some alternative embodiments, oxidases are used in combination with oxygen. Both types of enzymes are used for "solution bleaching" (i.e., to prevent transfer of a textile dye from a dyed fabric to another fabric when the fabrics are washed together in a wash liquor), preferably together with an enhancing agent (See e.g., WO94/12621 and WO95/01426). Suitable peroxidases/oxidases include, but are not limited to those of plant, bacterial or fungal origin. Chemically or genetically modified mutants are included in some embodiments. In some embodiments, the cleaning compositions of the present invention further comprise peroxidase and/or oxidase enzymes at a level from about 0.00001 to about 10% of additional peroxidase and/or oxidase by weight of the composition and the balance of cleaning adjunct materials by weight of composition. In some other embodiments of the present invention, the cleaning compositions of the present invention also comprise peroxidase and/or oxidase enzymes at a level of about 0.0001 to about 10%, about 0.001 to about 5%, about 0.001 to about 2%, about 0.005 to about 0.5% peroxidase and/or oxidase enzymes by weight of the composition.

**[00118]** In some embodiments, additional enzymes find use, including but not limited to perhydrolases (See e.g., WO2005056782, WO2007106293, WO2008063400, WO2008106214,

and WO2008106215). In addition, in some embodiments, mixtures of the above mentioned enzymes are encompassed herein, in particular one or more additional protease, amylase, lipase, mannanase, and/or at least one cellulase. Indeed, it is contemplated that various mixtures of these enzymes will find use in the present invention. It is also contemplated that the varying levels of the serine protease polypeptide (s) and one or more additional enzymes may both independently range to about 10%, the balance of the cleaning composition being cleaning adjunct materials. The specific selection of cleaning adjunct materials are readily made by considering the surface, item, or fabric to be cleaned, and the desired form of the composition for the cleaning conditions during use (e.g., through the wash detergent use).

**[00119]** In some embodiments, an effective amount of one or more serine protease polypeptide (s) provided herein is included in compositions useful for cleaning a variety of surfaces in need of proteinaceous stain removal. Such cleaning compositions include cleaning compositions for such applications as cleaning hard surfaces, fabrics, and dishes. Indeed, in some embodiments, the present invention provides fabric cleaning compositions, while in other embodiments the present invention provides non-fabric cleaning compositions. Notably, the present invention also provides cleaning compositions suitable for personal care, including oral care (including dentrifices, toothpastes, mouthwashes, etc., as well as denture cleaning compositions), skin, and hair cleaning compositions. It is intended that the present invention encompass detergent compositions in any form (i.e., liquid, granular, bar, semi-solid, gels, emulsions, tablets, capsules, etc.).

**[00120]** By way of example, several cleaning compositions wherein the serine protease polypeptides of the present invention find use are described in greater detail below. In some embodiments in which the cleaning compositions of the present invention are formulated as compositions suitable for use in laundry machine washing method(s), the compositions of the present invention preferably contain at least one surfactant and at least one builder compound, as well as one or more cleaning adjunct materials preferably selected from organic polymeric compounds, bleaching agents, additional enzymes, suds suppressors, dispersants, lime-soap dispersants, soil suspension and anti-redeposition agents and corrosion inhibitors. In some embodiments, laundry compositions also contain softening agents (i.e., as additional cleaning adjunct materials). The compositions of the present invention also find use in detergent additive products in solid or liquid form. Such additive products are intended to supplement and/or boost

the performance of conventional detergent compositions and can be added at any stage of the cleaning process. In some embodiments, the density of the laundry detergent compositions herein ranges from about 400 to about 1200 g/liter, while in other embodiments it ranges from about 500 to about 950 g/liter of composition measured at 20°C.

**[00121]** In embodiments formulated as compositions for use in manual dishwashing methods, the compositions of the invention preferably contain at least one surfactant and preferably at least one additional cleaning adjunct material selected from organic polymeric compounds, suds enhancing agents, group II metal ions, solvents, hydrotropes and additional enzymes.

**[00122]** In some embodiments, various cleaning compositions such as those provided in US 6,605,458, find use with the serine protease polypeptides of the present invention. Thus, in some embodiments, the compositions comprising at least one serine protease polypeptide of the present invention is a compact granular fabric cleaning composition, while in other embodiments, the composition is a granular fabric cleaning composition useful in the laundering of colored fabrics, in further embodiments, the composition is a granular fabric cleaning composition which provides softening through the wash capacity, in additional embodiments, the composition is a heavy duty liquid fabric cleaning composition. In some embodiments, the compositions comprising at least one serine protease polypeptide of the present invention are fabric cleaning compositions such as those described in US 6,610,642 and 6,376,450. In addition, the serine protease polypeptides of the present invention find use in granular laundry detergent compositions of particular utility under European or Japanese washing conditions (See e.g., US 6,610,642).

**[00123]** In some alternative embodiments, the present invention provides hard surface cleaning compositions comprising at least one serine protease polypeptide provided herein. Thus, in some embodiments, the compositions comprising at least one serine protease polypeptide of the present invention is a hard surface cleaning composition such as those described in US 6,610,642; 6,376,450; and 6,376,450.

**[00124]** In yet further embodiments, the present invention provides dishwashing compositions comprising at least one serine protease polypeptide provided herein. Thus, in some embodiments, the compositions comprising at least one serine protease polypeptide of the present invention is a hard surface cleaning composition such as those in US 6,610,642 and 6,376,450. In some still further embodiments, the present invention provides dishwashing

compositions comprising at least one serine protease polypeptide provided herein. In some further embodiments, the compositions comprising at least one serine protease polypeptide of the present invention comprise oral care compositions such as those in US 6,376,450, and 6,376,450. The formulations and descriptions of the compounds and cleaning adjunct materials contained in the aforementioned US 6,376,450; 6,605,458; 6,605,458; and 6,610,642 find use with the serine protease polypeptides provided herein.

**[00125]** The cleaning compositions of the present invention are formulated into any suitable form and prepared by any process chosen by the formulator, non-limiting examples of which are described in US 5,879,584; 5,691,297; 5,574,005; 5,569,645; 5,565,422; 5,516,448; 5,489,392; and 5,486,303, all of which are incorporated herein by reference. When a low pH cleaning composition is desired, the pH of such composition is adjusted via the addition of a material such as monoethanolamine or an acidic material such as HCl.

**[00126]** In some embodiments, the cleaning compositions according to the present invention comprise an acidifying particle or an amino carboxylic builder. Examples of an amino carboxylic builder include aminocarboxylic acids, salts and derivatives thereof. In some embodiment, the amino carboxylic builder is an aminopolycarboxylic builder, such as glycine-N,N-diacetic acid or derivative of general formula  $MOOC-CHR-N(CH_2COOM)_2$  where R is  $C_{1-12}$ alkyl and M is alkali metal. In some embodiments, the amino carboxylic builder can be methylglycine diacetic acid (MGDA), GLDA (glutamic-N,N-diacetic acid), iminodisuccinic acid (IDS), carboxymethyl inulin and salts and derivatives thereof, aspartic acid-N-monoacetic acid (ASMA), aspartic acid-N,N-diacetic acid (ASDA), aspartic acid-N-monopropionic acid (ASMP), iminodisuccinic acid (IDA), N-(2-sulfomethyl) aspartic acid (SMAS), N-(2-sulfoethyl)aspartic acid (SEAS), N-(2-sulfomethyl)glutamic acid (SMGL), N-(2-sulfoethyl) glutamic acid (SEGL), IDS (iminodiacetic acid) and salts and derivatives thereof such as N-methyliminodiacetic acid (MIDA), alpha-alanine-N,N-diacetic acid (alpha-ALDA), serine-N,N-diacetic acid (SEDA), isoserine-N,N-diacetic acid (ISDA), phenylalanine-N,N-diacetic acid (PHDA), anthranilic acid-N,N-diacetic acid (ANDA), sulfanilic acid-N,N-diacetic acid (SLDA), taurine-N,N-diacetic acid (TUDA) and sulfomethyl-N,N-diacetic acid (SMDA) and alkali metal salts and derivative thereof. In some embodiments, the acidifying particle has a weight geometric mean particle size of from about 400 $\mu$  to about 1200 $\mu$  and a bulk density of at least 550 g/L. In some embodiments, the acidifying particle comprises at least about 5% of the builder.

**[00127]** In some embodiments, the acidifying particle can comprise any acid, including organic acids and mineral acids. Organic acids can have one or two carboxyls and in some instances up to 15 carbons, especially up to 10 carbons, such as formic, acetic, propionic, capric, oxalic, succinic, adipic, maleic, fumaric, sebacic, malic, lactic, glycolic, tartaric and glyoxylic acids. In some embodiments, the acid is citric acid. Mineral acids include hydrochloric and sulphuric acid. In some instances, the acidifying particle of the invention is a highly active particle comprising a high level of amino carboxylic builder. Sulphuric acid has been found to further contribute to the stability of the final particle.

**[00128]** In some embodiments, the cleaning compositions according to the present invention comprise at least one surfactant and/or a surfactant system wherein the surfactant is selected from nonionic surfactants, anionic surfactants, cationic surfactants, ampholytic surfactants, zwitterionic surfactants, semi-polar nonionic surfactants and mixtures thereof. In some embodiments, the surfactant is present at a level of from about 0.1 to about 60%, while in alternative embodiments the level is from about 1 to about 50%, while in still further embodiments the level is from about 5 to about 40%, by weight of the cleaning composition.

**[00129]** In some embodiments, the cleaning compositions of the present invention comprise one or more detergent builders or builder systems. In some embodiments incorporating at least one builder, the cleaning compositions comprise at least about 1%, from about 3 to about 60% or even from about 5 to about 40% builder by weight of the cleaning composition. Builders include, but are not limited to, the alkali metal; ammonium and alkanolammonium salts of polyphosphates; alkali metal silicates; alkaline earth and alkali metal carbonates, aluminosilicates; ether hydroxypolycarboxylates; copolymers of maleic anhydride with ethylene or vinyl methyl ether; 1,3,5-trihydroxybenzene-2, 4,6-trisulphonic acid; ammonium and substituted ammonium salts of polyacetic acids such as ethylenediamine tetraacetic acid and nitrilotriacetic acid; polycarboxylates such as mellitic acid, succinic acid, citric acid, oxydisuccinic acid, polymaleic acid, benzene 1,3,5-tricarboxylic acid, carboxymethyloxysuccinic acid; and soluble salts thereof. Indeed, it is contemplated that any suitable builder will find use in various embodiments of the present invention.

**[00130]** In some embodiments, the builders form water-soluble hardness ion complexes (e.g., sequestering builders), such as citrates and polyphosphates (e.g., sodium tripolyphosphate and sodium tripolyphosphate hexahydrate, potassium tripolyphosphate, and mixed sodium and

potassium tripolyphosphate, etc.). It is contemplated that any suitable builder will find use in the present invention, including those known in the art (See e.g., EP2100949).

**[00131]** In some embodiments, builders for use herein include phosphate builders and non-phosphate builders. In some embodiments, the builder is a phosphate builder. In some embodiments, the builder is a non-phosphate builder. If present, builders are used in a level of from 0.1 to 80%, or from 5 to 60%, or from 10 to 50% by weight of the composition. In some embodiments the product comprises a mixture of phosphate and non-phosphate builders. Suitable phosphate builders include mono-phosphates, di-phosphates, tri-polyphosphates or oligomeric-poylphosphates, including the alkali metal salts of these compounds, including the sodium salts. In some embodiments, a builder can be sodium tripolyphosphate (STPP). Additionally, the composition can comprise carbonate and/or citrate, preferably citrate that helps to achieve a neutral pH composition of the invention. Other suitable non-phosphate builders include homopolymers and copolymers of polycarboxylic acids and their partially or completely neutralized salts, monomeric polycarboxylic acids and hydroxycarboxylic acids and their salts. In some embodiments, salts of the above mentioned compounds include the ammonium and/or alkali metal salts, i.e. the lithium, sodium, and potassium salts, including sodium salts. Suitable polycarboxylic acids include acyclic, alicyclic, hetero-cyclic and aromatic carboxylic acids, wherein in some embodiments, they can contain at least two carboxyl groups which are in each case separated from one another by, in some instances, no more than two carbon atoms.

**[00132]** In some embodiments, the cleaning compositions of the present invention contain at least one chelating agent. Suitable chelating agents include, but are not limited to copper, iron and/or manganese chelating agents and mixtures thereof. In embodiments in which at least one chelating agent is used, the cleaning compositions of the present invention comprise from about 0.1 to about 15% or even from about 3.0 to about 10% chelating agent by weight of the subject cleaning composition.

**[00133]** In some still further embodiments, the cleaning compositions provided herein contain at least one deposition aid. Suitable deposition aids include, but are not limited to, polyethylene glycol; polypropylene glycol; polycarboxylate; soil release polymers such as polytelephthalic acid; clays such as kaolinite, montmorillonite, atapulgitite, illite, bentonite, and halloysite; and mixtures thereof.

**[00134]** As indicated herein, in some embodiments, anti-redeposition agents find use in some



embodiments of the present invention. In some embodiments, non-ionic surfactants find use. For example, in automatic dishwashing embodiments, non-ionic surfactants find use for surface modification purposes, in particular for sheeting, to avoid filming and spotting and to improve shine. These non-ionic surfactants also find use in preventing the re-deposition of soils. In some embodiments, the anti-redeposition agent is a non-ionic surfactant as known in the art (See e.g., EP2100949). In some embodiments, the non-ionic surfactant can be ethoxylated nonionic surfactants, epoxy-capped poly(oxyalkylated) alcohols and amine oxides surfactants.

**[00135]** In some embodiments, the cleaning compositions of the present invention include one or more dye transfer inhibiting agents. Suitable polymeric dye transfer inhibiting agents include, but are not limited to, polyvinylpyrrolidone polymers, polyamine N-oxide polymers, copolymers of N-vinylpyrrolidone and N-vinylimidazole, polyvinylloxazolidones and polyvinylimidazoles or mixtures thereof. In embodiments in which at least one dye transfer inhibiting agent is used, the cleaning compositions of the present invention comprise from about 0.0001 to about 10%, from about 0.01 to about 5%, or even from about 0.1 to about 3% by weight of the cleaning composition.

**[00136]** In some embodiments, silicates are included within the compositions of the present invention. In some such embodiments, sodium silicates (e.g., sodium disilicate, sodium metasilicate, and crystalline phyllosilicates) find use. In some embodiments, silicates are present at a level of from about 1 to about 20%. In some embodiments, silicates are present at a level of from about 5 to about 15% by weight of the composition.

**[00137]** In some still additional embodiments, the cleaning compositions of the present invention also contain dispersants. Suitable water-soluble organic materials include, but are not limited to the homo- or co-polymeric acids or their salts, in which the polycarboxylic acid comprises at least two carboxyl radicals separated from each other by not more than two carbon atoms.

**[00138]** In some further embodiments, the enzymes used in the cleaning compositions are stabilized by any suitable technique. In some embodiments, the enzymes employed herein are stabilized by the presence of water-soluble sources of calcium and/or magnesium ions in the finished compositions that provide such ions to the enzymes. In some embodiments, the enzyme stabilizers include oligosaccharides, polysaccharides, and inorganic divalent metal salts, including alkaline earth metals, such as calcium salts, such as calcium formate. It is

contemplated that various techniques for enzyme stabilization will find use in the present invention. For example, in some embodiments, the enzymes employed herein are stabilized by the presence of water-soluble sources of zinc (II), calcium (II) and/or magnesium (II) ions in the finished compositions that provide such ions to the enzymes, as well as other metal ions (e.g., barium (II), scandium (II), iron (II), manganese (II), aluminum (III), Tin (II), cobalt (II), copper (II), nickel (II), and oxovanadium (IV)). Chlorides and sulfates also find use in some embodiments of the present invention. Examples of suitable oligosaccharides and polysaccharides (e.g., dextrans) are known in the art (See e.g., WO07/145964). In some embodiments, reversible protease inhibitors also find use, such as boron-containing compounds (e.g., borate, 4-formyl phenyl boronic acid) and/or a tripeptide aldehyde find use to further improve stability, as desired.

**[00139]** In some embodiments, bleach, bleach activators and/or bleach catalysts are present in the compositions of the present invention. In some embodiments, the cleaning compositions of the present invention comprise inorganic and/or organic bleaching compound(s). Inorganic bleaches include, but are not limited to perhydrate salts (e.g., perborate, percarbonate, perphosphate, persulfate, and persilicate salts). In some embodiments, inorganic perhydrate salts are alkali metal salts. In some embodiments, inorganic perhydrate salts are included as the crystalline solid, without additional protection, although in some other embodiments, the salt is coated. Any suitable salt known in the art finds use in the present invention (See e.g., EP2100949).

**[00140]** In some embodiments, bleach activators are used in the compositions of the present invention. Bleach activators are typically organic peracid precursors that enhance the bleaching action in the course of cleaning at temperatures of 60°C and below. Bleach activators suitable for use herein include compounds which, under perhydrolysis conditions, give aliphatic peroxy-carboxylic acids having preferably from about 1 to about 10 carbon atoms, in particular from about 2 to about 4 carbon atoms, and/or optionally substituted perbenzoic acid. Additional bleach activators are known in the art and find use in the present invention (See e.g., EP2100949).

**[00141]** In addition, in some embodiments and as further described herein, the cleaning compositions of the present invention further comprise at least one bleach catalyst. In some embodiments, the manganese triazacyclononane and related complexes find use, as well as

cobalt, copper, manganese, and iron complexes. Additional bleach catalysts find use in the present invention (See e.g., US 4,246,612; 5,227,084; 4,810,410; WO99/06521; and EP2100949).

**[00142]** In some embodiments, the cleaning compositions of the present invention contain one or more catalytic metal complexes. In some embodiments, a metal-containing bleach catalyst finds use. In some embodiments, the metal bleach catalyst comprises a catalyst system comprising a transition metal cation of defined bleach catalytic activity, (e.g., copper, iron, titanium, ruthenium, tungsten, molybdenum, or manganese cations), an auxiliary metal cation having little or no bleach catalytic activity (e.g., zinc or aluminum cations), and a sequesterant having defined stability constants for the catalytic and auxiliary metal cations, particularly ethylenediaminetetraacetic acid, ethylenediaminetetra (methylenephosphonic acid) and water-soluble salts thereof are used (See e.g., US 4,430,243). In some embodiments, the cleaning compositions of the present invention are catalyzed by means of a manganese compound. Such compounds and levels of use are well known in the art (See e.g., US 5,576,282). In additional embodiments, cobalt bleach catalysts find use in the cleaning compositions of the present invention. Various cobalt bleach catalysts are known in the art (See e.g., US 5,597,936 and 5,595,967) and are readily prepared by known procedures.

**[00143]** In some additional embodiments, the cleaning compositions of the present invention include a transition metal complex of a macropolycyclic rigid ligand (MRL). As a practical matter, and not by way of limitation, in some embodiments, the compositions and cleaning processes provided by the present invention are adjusted to provide on the order of at least one part per hundred million of the active MRL species in the aqueous washing medium, and in some embodiments, provide from about 0.005 to about 25 ppm, more preferably from about 0.05 to about 10 ppm, and most preferably from about 0.1 to about 5 ppm, of the MRL in the wash liquor.

**[00144]** In some embodiments, transition-metals in the instant transition-metal bleach catalyst include, but are not limited to manganese, iron and chromium. MRLs also include, but are not limited to special ultra-rigid ligands that are cross-bridged (e.g., 5,12-diethyl-1,5,8,12-tetraazabicyclo[6.6.2]hexadecane). Suitable transition metal MRLs are readily prepared by known procedures (See e.g., WO2000/32601, and US 6,225,464).

**[00145]** In some embodiments, the cleaning compositions of the present invention comprise metal care agents. Metal care agents find use in preventing and/or reducing the tarnishing,

corrosion, and/or oxidation of metals, including aluminum, stainless steel, and non-ferrous metals (e.g., silver and copper). Suitable metal care agents include those described in EP2100949, WO9426860 and WO94/26859). In some embodiments, the metal care agent is a zinc salt. In some further embodiments, the cleaning compositions of the present invention comprise from about 0.1 to about 5% by weight of one or more metal care agent.

**[00146]** In some embodiments, the cleaning composition is a high density liquid (HDL) composition having a variant serine protease polypeptide protease. The HDL liquid laundry detergent can comprise a deterative surfactant (10-40%) comprising anionic deterative surfactant selected from a group of linear or branched or random chain, substituted or unsubstituted alkyl sulphates, alkyl sulphonates, alkyl alkoxyated sulphate, alkyl phosphates, alkyl phosphonates, alkyl carboxylates, and/or mixtures thereof; and optionally non-ionic surfactant selected from a group of linear or branched or random chain, substituted or unsubstituted alkyl alkoxyated alcohol, for example, a C<sub>8</sub>-C<sub>18</sub>alkyl ethoxyated alcohol and/or C<sub>6</sub>-C<sub>12</sub>alkyl phenol alkoxyates, optionally wherein the weight ratio of anionic deterative surfactant (with a hydrophilic index (HIC) of from 6.0 to 9) to non-ionic deterative surfactant is greater than 1:1.

**[00147]** The composition can comprise optionally, a surfactancy boosting polymer consisting of amphiphilic alkoxyated grease cleaning polymers selected from a group of alkoxyated polymers having branched hydrophilic and hydrophobic properties, such as alkoxyated polyalkylenimines in the range of 0.05wt%-10wt% and/or random graft polymers typically comprising a hydrophilic backbone comprising monomers selected from the group consisting of: unsaturated C<sub>1</sub>-C<sub>6</sub>carboxylic acids, ethers, alcohols, aldehydes, ketones, esters, sugar units, alkoxy units, maleic anhydride, saturated polyalcohols such as glycerol, and mixtures thereof; and hydrophobic side chain(s) selected from the group consisting of: C<sub>4</sub>-C<sub>25</sub>alkyl group, polypropylene, polybutylene, vinyl ester of a saturated C<sub>2</sub>-C<sub>6</sub>mono-carboxylic acid, C<sub>1</sub>-C<sub>6</sub>alkyl ester of acrylic or methacrylic acid, and mixtures thereof.

**[00148]** The composition can comprise additional polymers such as soil release polymers including, for example, anionically end-capped polyesters, for example SRP1; polymers comprising at least one monomer unit selected from saccharide, dicarboxylic acid, polyol and combinations thereof, in random or block configuration; ethylene terephthalate-based polymers and co-polymers thereof in random or block configuration, for example, Repel-o-tex SF, SF-2 and SRP6, Texcare SRA100, SRA300, SRN100, SRN170, SRN240, SRN300 and SRN325,

Marloquest SL; anti-redeposition polymers (0.1 wt% to 10wt%, including, for example, carboxylate polymers, such as polymers comprising at least one monomer selected from acrylic acid, maleic acid (or maleic anhydride), fumaric acid, itaconic acid, aconitic acid, mesaconic acid, citraconic acid, methylenemalononic acid, and any mixture thereof; vinylpyrrolidone homopolymer; and/or polyethylene glycol with a molecular weight in the range of from 500 to 100,000 Da); cellulosic polymer (including, for example, alkyl cellulose; alkyl alkoxyalkyl cellulose; carboxyalkyl cellulose; alkyl carboxyalkyl cellulose, examples of which include carboxymethyl cellulose, methyl cellulose, methyl hydroxyethyl cellulose, methyl carboxymethyl cellulose; and mixtures thereof); and polymeric carboxylate (such as, for example, maleate/acrylate random copolymer or polyacrylate homopolymer).

**[00149]** The composition can further comprise saturated or unsaturated fatty acid, preferably saturated or unsaturated C<sub>12</sub>-C<sub>24</sub>fatty acid (0 to 10 wt%); deposition aids (including, for example, polysaccharides, cellulosic polymers, polydiallyl dimethyl ammonium halides (DADMAC), and co-polymers of DADMAC with vinyl pyrrolidone, acrylamides, imidazoles, imidazolinium halides, and mixtures thereof, in random or block configuration; cationic guar gum; cationic cellulose such as cationic hydroxyethyl cellulose; cationic starch; cationic polyacrylamides; and mixtures thereof.

**[00150]** The composition can further comprise dye transfer inhibiting agents examples of which include manganese phthalocyanine, peroxidases, polyvinylpyrrolidone polymers, polyamine N-oxide polymers, copolymers of N-vinylpyrrolidone and N-vinylimidazole, polyvinylloxazolidones and polyvinylimidazoles and/or mixtures thereof; chelating agents examples of which include ethylene-diamine-tetraacetic acid (EDTA); diethylene triamine penta methylene phosphonic acid (DTPMP); hydroxy-ethane diphosphonic acid (HEDP); ethylenediamine N,N'-disuccinic acid (EDDS); methyl glycine diacetic acid (MGDA); diethylene triamine penta acetic acid (DTPA); propylene diamine tetracetic acid (PDT A); 2-hydroxypyridine-N-oxide (HPNO); or methyl glycine diacetic acid (MGDA); glutamic acid N,N'-diacetic acid (N,N-dicarboxymethyl glutamic acid tetrasodium salt (GLDA); nitrilotriacetic acid (NTA); 4,5-dihydroxy-m-benzenedisulfonic acid; citric acid and any salts thereof; N-hydroxyethylethylenediaminetri-acetic acid (HEDTA), triethylenetetraaminehexaacetic acid (TTHA), N-hydroxyethyliminodiacetic acid (HEIDA), dihydroxyethylglycine (DHEG), ethylenediaminetetrapropionic acid (EDTP), and derivatives thereof.

**[00151]** The composition may comprise an enzyme stabilizer (examples of which include polyols such as propylene glycol or glycerol, sugar or sugar alcohol, lactic acid, reversible protease inhibitor, boric acid, or a boric acid derivative, e.g., an aromatic borate ester, or a phenyl boronic acid derivative such as 4-formylphenyl boronic acid).

**[00152]** The composition can further comprise silicone or fatty-acid based suds suppressors; heuing dyes, calcium and magnesium cations, visual signaling ingredients, anti-foam (0.001 to about 4.0 wt%), and/or structurant/thickener (0.01 to 5 wt%, selected from the group consisting of diglycerides, triglycerides, ethylene glycol distearate, microcrystalline cellulose, cellulose based materials, microfiber cellulose, biopolymers, xanthan gum, gellan gum, and mixtures thereof).

**[00153]** Suitable deterative surfactants also include cationic deterative surfactants (selected from a group of alkyl pyridinium compounds, alkyl quarternary ammonium compounds, alkyl quarternary phosphonium compounds, alkyl ternary sulphonium compounds, and/or mixtures thereof); zwitterionic and/or amphoteric deterative surfactants (selected from a group of alkanolamine sulfo-betaines); ampholytic surfactants; semi-polar non-ionic surfactants; and mixtures thereof.

**[00154]** The composition can be any liquid form, for example a liquid or gel form, or any combination thereof. The composition may be in any unit dose form, for example a pouch.

**[00155]** In some embodiments, the cleaning composition is a high density powder (HDD) composition having a variant serine protease polypeptide protease. The HDD powder laundry detergent can comprise a deterative surfactant including anionic deterative surfactants (selected from a group of linear or branched or random chain, substituted or unsubstituted alkyl sulphates, alkyl sulphonates, alkyl alkoxyated sulphate, alkyl phosphates, alkyl phosphonates, alkyl carboxylates and/or mixtures thereof), non-ionic deterative surfactant (selected from a group of linear or branched or random chain, substituted or unsubstituted C<sub>8</sub>-C<sub>18</sub> alkyl ethoxylates, and/or C<sub>6</sub>-C<sub>12</sub> alkyl phenol alkoxyates), cationic deterative surfactants (selected from a group of alkyl pyridinium compounds, alkyl quaternary ammonium compounds, alkyl quaternary phosphonium compounds, alkyl ternary sulphonium compounds, and mixtures thereof), zwitterionic and/or amphoteric deterative surfactants (selected from a group of alkanolamine sulfo-betaines); ampholytic surfactants; semi-polar non-ionic surfactants and mixtures thereof; builders (phosphate free builders [for example zeolite builders examples of which include zeolite A,

zeolite X, zeolite P and zeolite MAP in the range of 0 to less than 10 wt%]; phosphate builders [examples of which include sodium tri-polyphosphate in the range of 0wt% to less than 10wt%]; citric acid, citrate salts and nitrilotriacetic acid or salt thereof in the range of less than 15 wt%); silicate salt (sodium or potassium silicate or sodium meta-silicate in the range of 0wt% to less than 10wt%, or layered silicate (SKS-6)); carbonate salt (sodium carbonate and/or sodium bicarbonate in the range of 0 to less than 10 wt%); and bleaching agents (photobleaches, examples of which include sulfonated zinc phthalocyanines, sulfonated aluminum phthalocyanines, xanthenes dyes, and mixtures thereof; hydrophobic or hydrophilic bleach activators (examples of which include dodecanoyl oxybenzene sulfonate, decanoyl oxybenzene sulfonate, decanoyl oxybenzoic acid or salts thereof, 3,5,5-trimethyl hexanoyl oxybenzene sulfonate, tetraacetyl ethylene diamine-TAED, and nonanoyloxybenzene sulfonate-NOBS, nitrile quats, and mixtures thereof; hydrogen peroxide; sources of hydrogen peroxide (inorganic perhydrate salts examples of which include mono or tetra hydrate sodium salt of perborate, percarbonate, persulfate, perphosphate, or persilicate); preformed hydrophilic and/or hydrophobic peracids (selected from a group consisting of percarboxylic acids and salts, percarbonic acids and salts, perimidic acids and salts, peroxymonosulfuric acids and salts) & mixtures thereof and/or bleach catalyst (such as imine bleach boosters examples of which include iminium cations and polyions; iminium zwitterions; modified amines; modified amine oxides; N-sulphonyl imines; N-phosphonyl imines; N-acyl imines; thiadiazole dioxides; perfluoroimines; cyclic sugar ketones and mixtures thereof; metal-containing bleach catalyst for example copper, iron, titanium, ruthenium, tungsten, molybdenum, or manganese cations along with an auxiliary metal cations such as zinc or aluminum and a sequestrate such as ethylenediaminetetraacetic acid, ethylenediaminetetra(methylenephosphonic acid) and water-soluble salts thereof ).

**[00156]** The composition can further comprise additional detergent ingredients including perfume microcapsules, starch encapsulated perfume accord, hueing agents, additional polymers including fabric integrity and cationic polymers, dye lock ingredients, fabric-softening agents, brighteners (for example C.I. Fluorescent brighteners), flocculating agents, chelating agents, alkoxyated polyamines, fabric deposition aids, and/or cyclodextrin.

**[00157]** In some embodiments, the cleaning composition is an ADW detergent composition having a serine protease of the present invention. The ADW detergent composition can

comprise two or more non-ionic surfactants selected from a group of ethoxylated non-ionic surfactants, alcohol alkoxylated surfactants, epoxy-capped poly(oxyalkylated) alcohols, or amine oxide surfactants present in amounts from 0 to 10% by weight; builders in the range of 5-60% comprising either phosphate (mono-phosphates, di-phosphates, tri-polyphosphates or oligomeric-polyphosphates, preferred sodium tripolyphosphate-STPP or phosphate-free builders [amino acid based compounds, examples of which include MGDA (methyl-glycine-diacetic acid), and salts and derivatives thereof, GLDA (glutamic-N,N-diacetic acid) and salts and derivatives thereof, IDS (iminodisuccinic acid) and salts and derivatives thereof, carboxy methyl inulin and salts and derivatives thereof and mixtures thereof, nitrilotriacetic acid (NTA), diethylene triamine penta acetic acid (DTPA), B-alaninediacetic acid (B-ADA) and their salts], homopolymers and copolymers of poly-carboxylic acids and their partially or completely neutralized salts, monomeric polycarboxylic acids and hydroxycarboxylic acids and their salts in the range of 0.5 to 50% by weight; sulfonated/carboxylated polymers (provide dimensional stability to the product) in the range of about 0.1 to about 50% by weight; drying aids in the range of about 0.1 to about 10% by weight (selected from polyesters, especially anionic polyesters optionally together with further monomers with 3 to 6 functionalities which are conducive to polycondensation, specifically acid, alcohol or ester functionalities, polycarbonate-, polyurethane- and/or polyurea-polyorganosiloxane compounds or precursor compounds thereof of the reactive cyclic carbonate and urea type); silicates in the range from about 1 to about 20% by weight (sodium or potassium silicates for example sodium disilicate, sodium meta-silicate and crystalline phyllosilicates); bleach-inorganic (for example perhydrate salts such as perborate, percarbonate, perphosphate, persulfate and persilicate salts) and organic (for example organic peroxyacids including diacyl and tetraacylperoxides, especially diperoxydodecanedioic acid, diperoxytetradecanedioic acid, and diperoxyhexadecanedioic acid); bleach activators- organic peracid precursors in the range from about 0.1 to about 10% by weight; bleach catalysts (selected from manganese triazacyclononane and related complexes, Co, Cu, Mn and Fe bispyridylamine and related complexes, and pentamine acetate cobalt(III) and related complexes); metal care agents in the range from about 0.1 to 5% by weight (selected from benzotriazoles, metal salts and complexes, and/or silicates); enzymes in the range from about 0.01 to 5.0mg of active enzyme per gram of automatic dishwashing detergent composition (acyl transferases, alpha-amylases, beta-amylases, alpha-galactosidases, arabinosidases, aryl esterases, beta-galactosidases,



carrageenases, catalases, cellobiohydrolases, cellulases, chondroitinases, cutinases, endo-beta-1, 4-glucanases, endo-beta-mannanases, esterases, exo-mannanases, galactanases, glucoamylases, hemicellulases, hyaluronidases, keratinases, laccases, lactases, ligninases, lipases, lipoxygenases, mannanases, oxidases, pectate lyases, pectin acetyl esterases, pectinases, pentosanases, peroxidases, phenoloxidases, phosphatases, phospholipases, phytases, polygalacturonases, proteases, pullulanases, reductases, rhamnogalacturonases, beta-glucanases, tannases, transglutaminases, xylan acetyl-esterases, xylanases, xyloglucanases, and xylosidases, and any mixture thereof); and enzyme stabilizer components (selected from oligosaccharides, polysaccharides and inorganic divalent metal salts).

**[00158]** In some embodiments, the cleaning composition is borate-free. In some embodiments, the cleaning composition is phosphate-free.

**[00159]** Representative detergent formulations that beneficially include a serine protease polypeptide of the present invention include the detergent formulations found in WO2013063460, pages 78-152, and in particular the tables of pages 94 to 152 are hereby incorporated by reference. The serine proteases are normally incorporated into the detergent composition at a level of from 0.00001 to 10% of enzyme protein by weight of the composition. In some embodiments, the detergent composition comprises more than 0.0001%, 0.001%, 0.01%, or 0.1% of the serine protease by weight of the composition. In some embodiments, the detergent composition comprises less than 1%, 0.1%, 0.01%, or 0.001% of the serine protease by weight of the composition.

**[00160]** Also provided are compositions and methods of treating fabrics (*e.g.*, to desize a textile) using a serine protease polypeptide of the present invention. Fabric-treating methods are well known in the art (see, *e.g.*, US 6,077,316). For example, the feel and appearance of a fabric can be improved by a method comprising contacting the fabric with a serine protease in a solution. The fabric can be treated with the solution under pressure.

**[00161]** A serine protease of the present invention can be applied during or after the weaving of a textile, or during the desizing stage, or one or more additional fabric processing steps. During the weaving of textiles, the threads are exposed to considerable mechanical strain. Prior to weaving on mechanical looms, warp yarns are often coated with sizing starch or starch derivatives to increase their tensile strength and to prevent breaking. A serine protease of the present invention can be applied during or after the weaving to remove the sizing starch or starch

derivatives. After weaving, the serine protease can be used to remove the size coating before further processing the fabric to ensure a homogeneous and wash-proof result.

**[00162]** A serine protease of the present invention can be used alone or with other desizing chemical reagents and/or desizing enzymes to desize fabrics, including cotton-containing fabrics, as detergent additives, *e.g.*, in aqueous compositions. An amylase also can be used in compositions and methods for producing a stonewashed look on indigo-dyed denim fabric and garments. For the manufacture of clothes, the fabric can be cut and sewn into clothes or garments, which are afterwards finished. In particular, for the manufacture of denim jeans, different enzymatic finishing methods have been developed. The finishing of denim garment normally is initiated with an enzymatic desizing step, during which garments are subjected to the action of proteolytic enzymes to provide softness to the fabric and make the cotton more accessible to the subsequent enzymatic finishing steps. The serine protease can be used in methods of finishing denim garments (*e.g.*, a “bio-stoning process”), enzymatic desizing and providing softness to fabrics, and/or finishing process.

**[00163]** The serine protease polypeptides described herein find further use in the enzyme aided removal of proteins from animals and their subsequent degradation or disposal, such as feathers, skin, hair, hide, and the like. In some instances, immersion of the animal carcass in a solution comprising a serine protease polypeptide of the present invention can act to protect the skin from damage in comparison to the traditional immersion in scalding water or the defeathering process. In one embodiment, feathers can be sprayed with an isolated serine protease polypeptide of the present invention under conditions suitable for digesting or initiating degradation of the plumage. In some embodiments, a serine protease of the present invention can be used, as above, in combination with an oxidizing agent.

**[00164]** In some embodiments, removal of the oil or fat associated with raw feathers is assisted by using a serine protease polypeptide of the present invention. In some embodiments, the serine protease polypeptides are used in compositions for cleaning the feathers as well as to sanitize and partially dehydrate the fibers. In yet other embodiments, the disclosed serine protease polypeptides find use in recovering protein from plumage. In some other embodiments, the serine protease polypeptides are applied in a wash solution in combination with 95% ethanol or other polar organic solvent with or without a surfactant at about 0.5% (v/v).

**[00165]** In a further aspect of the invention, the serine protease polypeptides of the present invention can be used as a component of an animal feed composition, animal feed additive and/or pet food comprising a serine protease and variants thereof. The present invention further relates to a method for preparing such an animal feed composition, animal feed additive composition and/or pet food comprising mixing the serine protease polypeptide with one or more animal feed ingredients and/or animal feed additive ingredients and/or pet food ingredients. Furthermore, the present invention relates to the use of the serine protease polypeptide in the preparation of an animal feed composition and/or animal feed additive composition and/or pet food.

**[00166]** The term “animal” includes all non-ruminant and ruminant animals. In a particular embodiment, the animal is a non-ruminant animal, such as a horse and a mono-gastric animal. Examples of mono-gastric animals include, but are not limited to, pigs and swine, such as piglets, growing pigs, sows; poultry such as turkeys, ducks, chicken, broiler chicks, layers; fish such as salmon, trout, tilapia, catfish and carps; and crustaceans such as shrimps and prawns. In a further embodiment the animal is a ruminant animal including, but not limited to, cattle, young calves, goats, sheep, giraffes, bison, moose, elk, yaks, water buffalo, deer, camels, alpacas, llamas, antelope, pronghorn and nilgai.

**[00167]** In the present context, it is intended that the term “pet food” is understood to mean a food for a household animal such as, but not limited to, dogs, cats, gerbils, hamsters, chinchillas, fancy rats, guinea pigs; avian pets, such as canaries, parakeets, and parrots; reptile pets, such as turtles, lizards and snakes; and aquatic pets, such as tropical fish and frogs.

**[00168]** The terms “animal feed composition,” “feedstuff” and “fodder” are used interchangeably and can comprise one or more feed materials selected from the group comprising a) cereals, such as small grains (e.g., wheat, barley, rye, oats and combinations thereof) and/or large grains such as maize or sorghum; b) by products from cereals, such as corn gluten meal, Distillers Dried Grain Solubles (DDGS) (particularly corn based Distillers Dried Grain Solubles (cDDGS), wheat bran, wheat middlings, wheat shorts, rice bran, rice hulls, oat hulls, palm kernel, and citrus pulp; c) protein obtained from sources such as soya, sunflower, peanut, lupin, peas, fava beans, cotton, canola, fish meal, dried plasma protein, meat and bone meal, potato protein, whey, copra, sesame; d) oils and fats obtained from vegetable and animal sources; and e) minerals and vitamins.

**[00169]** The protease polypeptides described herein find further use in the enzyme aided bleaching of paper pulps such as chemical pulps, semi-chemical pulps, kraft pulps, mechanical pulps or pulps prepared by the sulfite method. In general terms, paper pulps are incubated with a protease polypeptide of the present invention under conditions suitable for bleaching the paper pulp.

**[00170]** In some embodiments, the pulps are chlorine free pulps bleached with oxygen, ozone, peroxide or peroxyacids. In some embodiments, the protease polypeptides are used in enzyme aided bleaching of pulps produced by modified or continuous pulping methods that exhibit low lignin contents. In some other embodiments, the protease polypeptides are applied alone or preferably in combination with xylanase and/or endoglucanase and/or alpha-galactosidase and/or cellobiohydrolase enzymes.

**[00171]** The protease polypeptides described herein find further use in the enzyme aided removal of proteins from animals and their subsequent degradation or disposal, such as feathers, skin, hair, hide, and the like. In some instances, immersion of the animal carcass in a solution comprising a protease polypeptide of the present invention can act to protect the skin from damage in comparison to the traditional immersion in scalding water or the defeathering process. In one embodiment, feathers can be sprayed with an isolated protease polypeptide of the present invention under conditions suitable for digesting or initiating degradation of the plumage. In some embodiments, a protease of the present invention can be used, as above, in combination with an oxidizing agent.

**[00172]** In some embodiments, removal of the oil or fat associated with raw feathers is assisted by using a protease polypeptide of the present invention. In some embodiments, the protease polypeptides are used in compositions for cleaning the feathers as well as to sanitize and partially dehydrate the fibers. In some other embodiments, the protease polypeptides are applied in a wash solution in combination with 95% ethanol or other polar organic solvent with or without a surfactant at about 0.5% (v/v). In yet other embodiments, the disclosed protease polypeptides find use in recovering protein from plumage. The disclosed protease polypeptides may be used alone or in combination in suitable feather processing and proteolytic methods, such as those disclosed in PCT/EP2013/065362, PCT/EP2013/065363, and PCT/EP2013/065364, which are hereby incorporated by reference. In some embodiments, the recovered protein can be

subsequently used in animal or fish feed.

### EXAMPLES

[00173] The following examples are provided to demonstrate and illustrate certain preferred embodiments and aspects of the present disclosure and should not be construed as limiting. In the experimental disclosure which follows, the following abbreviations apply: ADW (automatic dish washing); BMI (blood/milk/ink); BSA (bovine serum albumin); CAPS (*N*-cyclohexyl-3-aminopropanesulfonic acid); CHES (*N*-cyclohexyl-2-aminoethanesulfonic acid); DMC (dimethyl casein); HDD (heavy duty dry/powder); HDL (heavy duty liquid); HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid); MTP (microtiter plate); ND (not done); OD (optical density); PCR (polymerase chain reaction); ppm (parts per million); QS (quantity sufficient); rpm (revolutions per minute); AAPF (succinyl-Ala-Ala-Pro-Phe-p-nitroanilide); TNBSA (2,4,6-trinitrobenzene sulfonic acid); v/v (volume to volume); and w/v (weight to volume).

#### EXAMPLE 1

##### Protein Determination Methods

###### Protein Determination by Stain Free Imager Criterion

[00174] Protein was quantified by the stain-free Imager Criterion method. The method is based on utilizing of stain-free precast PAGE gels, where the intensity of each band will depend on amount of tryptophan residues presented in the protein of interest. The Criterion™ TGX (Tris-Glycine extended) Stain-Free™ precast gels for PAGE include unique trihalo compounds. This allows rapid fluorescent detection of proteins with the Gel Doc™ EZ imaging system. The trihalo compounds react with tryptophan residues in a UV-induced reaction to produce fluorescence, which can be easily detected by the Gel Doc EZ imager within gels. Reagents used in the assay: Concentrated (10x) Laemmli Sample Buffer (Kem-En-Tec, Catalogue #42556); either 18 or 26-well Criterion TGX Stain-Free Precast gels (Bio-Rad, Catalogue #567-8124 and 567-8125, respectively); and protein markers "Precision Plus Protein Standards" (Bio-Rad, Catalogue #161-0363). The assay was carried on as follow: 25µl protein sample and 25µl 0.5M HCL was added to a 96-well PCR plate on ice for 10 min to inactivate the protease and prevent self-hydrolysis. 50µl of the acid protein mix was added to a 50µL sample buffer containing 0.385 mg DTT in the 96-well PCR plate. After that the chamber was filled by running buffer, and gel cassette was set. Then 10µL of each sample together with markers was load in each pocket and electrophoresis was started at 200 V for 35 min. Following electrophoresis, the gel was transferred to the Imager. Image Lab software was used for calculation of intensity of each

band. By knowing the protein amount and the tryptophan content of the standard sample, the calibration curve can be made. The amount of experimental sample can be determined by extrapolation of the band intensity and tryptophan numbers to protein concentration. The protein quantification method was employed to prepare samples of BspE04637 protease used for assays shown in subsequent Examples.

#### N and C-terminal amino acid determination

**[00175]** In preparation for sequence confirmation, a sample of isolated protein is subjected to a series of chemical treatments in a 10 kDa spinfilter. The sample is denatured and reduced by urea and DTT treatment. A guanidination step was performed to convert lysines to homoarginines to protect lysine side chains from acetylation. Acetylation reaction using iodoacetamide then modifies only the proteins' N-terminal residue. The sample is then mixed with a buffer containing  $^{18}\text{O}$  water and the enzymes trypsin and chymotrypsin are added for digestion. The resulting peptides will contain mixtures of  $^{18}\text{O}$  and  $^{16}\text{O}$ , except for the Carboxyl terminus which will retain the native  $^{16}\text{O}$ . The digestion products were separated and analyzed using a Proxeon nano-LC system followed by LTQ Orbitrap (Thermo Fisher) high resolution mass spectrometer and the amino acid sequence was deduced from the MS/MS fragment spectrum of the peptides, and the isotopic pattern of the peptides.

## EXAMPLE 2

### Discovery and Identification of serine protease BspE04637

**[00176]** *Bacillus sp.* GX6638 (ATCC 53278) was selected as a potential source for enzymes useful in industrial applications. To identify enzymes produced by *Bacillus sp.* GX6638 and the genes that encode these enzymes, the entire genome of *Bacillus sp.* GX6638 was sequenced using Illumina<sup>®</sup> sequencing by synthesis (SBS) technology. Genome sequencing and assembly of the sequence data was performed by BaseClear (Leiden, The Netherlands). Contigs were annotated by BioXpr (Namur, Belgium). One of genes identified this way in strain *Bacillus sp.* GX6638 encodes a protein that showed homology to serine proteases of various other bacteria. The sequence of this gene, *BspE04637.n*, is depicted in SEQ ID NO:1: ATGAGAAACTACTA ACCTTGTTAACGTTAAGCATTCTTGTATTTTCGATGGTAGTACCAATGTCGATTAGTGCTCAATCACAGGCAGAAAAACAGGAGTACCTTGTTCAATTTAATGACAAGGTAAT AAGGGGATATTAACGCATTTGGTGTAGATAACAGTGATGTTCTTCATACTTACAATCTACTACCTGTTAATCTAGTAAAAATGACAGAGCAGCAAGCGAAAGCATTACAAAA

TAATCCGCATATTAAGCGGTGGAACCTAACTTTGAGGCGCAAGCATTTCGCTCAAAC  
 AGTACCGTGGGGAGTTCCTCATGTTCAAGGTACTGATGCTCATGCAGCGGGGCATAC  
 TGGAAGTGGTGTAAAAGTAGCTATACTCGATACGGGAATTGATCGAAATCATGAAG  
 ATTTAAATGTAAGAGGAGGGCACTCTGTATTCACAGACAGTGCAAACAGGGATCCT  
 TATTACGATGGTAGTGGTCATGGCACGCACGTTGCTGGAACAGTAGCGGCCTTAAAT  
 AATAGTGTAGGTGTGTTAGGTGTAGCCTACAACGCGGAGCTATATGCGGTAAAAGT  
 ATTAATAACAGTGGAAGTGGTTCTTATGCGGGAATCGCTGAAGGAATTGAATGGG  
 CAGTACAAAACAATATGGACATTATAAATATGAGTTTAGGAGGTTTCGATGAGTTCTT  
 CTATTTTAGAAGAGTGGTGTAATATTGCCTACAATTCTGGTGTGCTTGTAGTGGCGG  
 CAGCTGGAAATAGTGGACGTACGAATGGCCGAGGGGATACTGTTGGTTATCCAGCA  
 AAATATGATTCTGTTATCGCAGTAGCAGCAGTTGATTCTAGCAACAATCGTGCAAGT  
 TTCTCTAGTACAGGACCTGCAGTGGAAATTGCTGCACCAGGTGTAAACATTCTTAGT  
 ACGACTCCTGGAAACAGCTATGCATCTTACAACGGAACCTTCGATGGCATCTCCACAT  
 GTAGCTGGTGTGGCAGCGCTTGTCTTGGCAGCGAATCCGAACCTTTCAAACGTGGAG  
 CTTCGTAATCGATTAAACGATACAGCGCAAACTTAGGAGATGCAAACCACTTTGG  
 AAACGGCCTAGTGCGTGCAGTTGATGCCATTAACGGCACTAGCTCCGGTGATAACG  
 GTGGCGGAAGCGAACCAACAAAACCAGGTAACGGCAAAGGAAACGGAAGAAAC.

[00177] The preproenzyme encoded by the *BspE04637.n* gene is depicted in SEQ ID NO:2. At the N-terminus, the protein has a signal peptide with a length of 25 amino acids as predicted by SignalP-NN (Emanuelsson *et al.*, Nature Protocols (2007)2:953-971). This signal peptide sequence is underlined and in bold in SEQ ID NO:2. The presence of a signal peptide indicates that this serine protease is a secreted enzyme. The enzyme has a pro sequence (shown in italics) which is predicted to be 72 amino acids. The sequence of the predicted, fully processed mature chain (*BspE04637*, 303 amino acids) is depicted in SEQ ID NO:3.

[00178] SEQ ID NO:2 sets forth the amino acid sequence of the serine protease precursor *BspE04637* (the signal peptide sequence is underlined and in bold, the pro sequence is shown in italics): **MRKLLTLLTSLVFSMVVPMSSISA***QSQAQEKQEYLVQFNDKVNKGILNAFGVDN*  
*SDVLHTYNLLPVNLVKMTEQQAKALQNNPHIKAVEPNFEAQAFQAQTVPWGVPHVQGTDAH*  
*AAGHTGSGVKVAILDTGIDRNHEDLNVRGGHSVFTDSANRDPYYDGSGHGTHVAGTV*  
*AALNNSVGLGVAYNAELYAVKVLNNSGSGSYAGIAEGIEWAVQNNMDIINMSLGS*  
*MSSSILEEWCNIAYNISGVLVVAAGNSGRTNGRGDTVGYPAKYDSVIAVAADVSSNNR*

ASFSSTGPAVEIAAPGVNILSTTPGNSYASYNGTSMASPHVAGVAALVLAANPNLSNVE  
LRNRLNDTAQNLGDANHFNGLVRAVDAINGTSSGDNNGGSEPTKPGNGKGNRN.

[00179] SEQ ID NO:3 sets forth the predicted amino acid sequence of the mature protease BspE04637: AQTVPWGVPHVQGTDAHAAGHTGSGVKVAILDGTIDRNHEDLNVRGGHS  
VFTDSANRDPYYDGSFGHGHVAGTVAALNNSVGVLGVAAYNAELYAVKVLNNSGSGSY  
AGIAEGIEWAVQNNMDIINMSLGGSMSSILEEWCNIAYN SGVLVVAAGNSGRTNGRG  
DTVGYPKYDSVIAVAVDSSNNRASFSSTGPAVEIAAPGVNILSTTPGNSYASYNGTS  
MASPHVAGVAALVLAANPNLSNVELRNRLNDTAQNLGDANHFNGLVRAVDAINGTS  
SGDNNGGSEPTKPGNGKGNRN.

### EXAMPLE 3

#### Discovery and Identification of serine protease SWT183\_1430046

[00180] *Bacillus sp.* SWT183 (Dupont Culture Collection) was also selected as a potential source for enzymes useful in industrial applications. To identify enzymes produced by *Bacillus sp.* SWT183 and the genes that encode these enzymes, the entire genome of *Bacillus sp.* SWT183 was sequenced using Illumina<sup>®</sup> sequencing by synthesis (SBS) technology. Genome sequencing, assembly of the sequence data, and annotation of the contigs was performed by BaseClear (Leiden, The Netherlands). One of genes identified this way in strain *Bacillus sp.* SWT183 encodes a protein that showed homology to BspE04637. The sequence of this gene, *SWT183\_1430046.n*, is depicted in SEQ ID NO:4.

[00181] SEQ ID NO:4 sets forth the nucleotide sequence of the *SWT183\_1430046.n* gene:  
ATGAGAAA ACTACTAACCTTGTTAACGTTAAGCATCCTTGTA TTTTCGATGTTAGTAC  
CAATGTCGATTAGTGCACAATCACAGGCAGAAAAACAGGAGTACCTTGTTCAATTTA  
ATGATCAGGTGAATAATGGAATATTAATGCATTTGGTG TAAAAGACAGTGATGTTC  
TTCATACTTACAATCTACTACCTGTCAATCTAGTAAAAATGACAGAGCAGCAAGCGA  
AAGCATTACAAAATAATCCACATATTAAGCGGTGGAACCTA ACTTTGAGGCGCAA  
GCATTCGCTCAAACCGTACCGTGGGGAGTTCCTCATGTTCAAGGTACCGATGCTCAC  
GCAGCGGGGCATACTGGAAGTGGTG TAAAAGTAGCTATACTTGATACGGGAATTGA  
TCGAAATCATGAAGATTTAAACGTAAGAGGAGGGCATTCTGTATTTACGGACAGTG  
CAAACAGCGATCCTTATTACGATGGTAGTGGTCATGGCACGCACGTTGCTGGAACAG  
TAGCGGCGTTAAATAATAGTGTAGGTGTGTTAGGTGTAGCCTACAATGCGGA ACTGT  
ATGCGGTAAAAGTATTAATAACAGTGGAAGTGGTTCTTACGCAGGAATTGCACAA



GGGATTGAATGGGCAGTTCAAAAACAATATGGACATTATTAATATGAGCTTAGGAGG  
 TTCGATGAGTTCTTCTATTTTAGAAGAATGGTGTAACATTGCTTACAATTCTGGTGTA  
 CTAGTAGTGGCAGCAGCTGGAAACAGTGGACGTACGAACGGTCGAGGAGATACTGT  
 CCGTTATCCAGCAAAATATGATTCTGTTATCGCAGTAGCAGCAGTCGATTCTAGCAA  
 TAACCGTGCAAGCTTCTCTAGTACAGGATCTGCAGTGGAGATTGCTGCACCAGGTGT  
 AACATTCTTAGCACAACTCCTGGAAACAGTTACGCATCTTACAACGGAACTTCGAT  
 GGCATCTCCACATGTAGCTGGTGTAGCAGCACTTGTATGGGCAGCAAATCCGAACCT  
 TTCAAACGTGGAGCTCCGTAATCGATTAAACGATACAGCGCAAACTTAGGTGATG  
 CAAACCACTTCGGACACGGCCTAGTCCGTGCAGTCGATGCCATTAACGGCACTAGCT  
 CCGGTGATAACGGTGGCGGTGACGATGGTGGAAAGCGGACCAACAAAACCAGGTAAC  
 GGCAAAGGAAACGGAAAAAACTAA.

[00182] The preproenzyme encoded by the *SWT183\_1430046.n* gene is depicted in SEQ ID NO:5. At the N-terminus, the protein has a signal peptide with a length of 25 amino acids as predicted by SignalP-NN (Emanuelsson *et al.*, Nature Protocols (2007)2:953-971). This signal peptide sequence is underlined and in bold in SEQ ID NO:5. The presence of a signal peptide indicates that this serine protease is a secreted enzyme. The enzyme has a pro sequence which is predicted to be 72 amino acids (shown in italics). The sequence of the predicted, fully processed mature chain (*SWT183\_1430046*, 307 amino acids) is depicted in SEQ ID NO:6.

[00183] SEQ ID NO:5 sets forth the amino acid sequence of the serine protease precursor *SWT183\_1430046* (the signal peptide sequence is underlined and in bold, the pro sequence is shown in italics): **MRKLLTLLTSLVFSMLVPMSSIS***QSQA EKQEYLVQFNDQVNNGILN AFGVKDS DVLHTYNLLPVNLVKMTEQQAKALQNNPHIKAVEPNFEAQAF AQTVPWGVPHV QGTD AHAAGHTGSGVKVAILDTGIDRNHEDLNVRGGH SVFTDSANSDPYD GSGHGTH VAGTVAALNNSVGLGVAYNAELYAVKVLNNSGSGSYAGIAQGI EWAVQNNMDIINM SLGGS MSSSILEEWCNIA YNSGVLVVAAGNSGRTN GRGDTVGYPAKYDSVIAVA AVD SSNNRASFSSTGSAVEIAAPGVN ILSTTPGNSYAS YNGTSMASPHVAGVAALVWA ANPN LSNVELRNRLNDTAQNLGDANHFHGLVRAVD AINGTSSGDNGGGDDGGSGPTKPGN GKGN GKN.*

[00184] SEQ ID NO:6 sets forth the predicted amino acid sequence of the mature protease *SWT183\_1430046*: AQTVPWGVPHVQGTD AHAAGHTGSGVKVAILDTGIDRNHEDLNVR GGH SVFTDSANSDPYD GSGHGTHVAGTVAALNNSVGLGVAYNAELYAVKVLNNSG

SGSYAGIAQGIEWAVQNNMDIINMSLGGSMSSSILEEWCNIAYNSGVLVVAAGNSGRT  
 NGRGDTVGYPAKYDSVIAVAADVSSNNRASFSSTGSAVEIAAPGVNILSTTPGNSYASY  
 NGTSMASPHVAGVAALVWAANPNLSNVELRNRLNDTAQNLGDANHFHGLVRAVDAI  
 NGTSSGDNGGGDDGGSGPTKPGNGKGNKGN.

#### EXAMPLE 4

##### Heterologous expression of BspE04637

[00185] BspE04637 protease was produced in *B. subtilis* using an expression cassette consisting of the *B. subtilis aprE* promoter, the *B. subtilis aprE* signal peptide sequence, the native BspE04637 protease pro-peptide, the mature BspE04637 protease and a BPN' terminator. This cassette was cloned into the pBN based replicating shuttle vector (Babe' et al. (1998), *Biotechnol. Appl. Biochem.* 27:117–124) and transformed into *B. subtilis* strain BG3594. A map of the pBN vector containing the BspE04637 gene (pBN-BspE04637) is shown in Figure 1.

[00186] To produce BspE04637, a *B. subtilis* transformant containing pBN-BspE04637 was cultured in 15 ml Falcon tubes for 16 hours in TSB (broth) with 10 ppm neomycin, and 300 µl of this pre-culture was added to a 500 mL flask filled with 30 mL of cultivation media (described below) supplemented with 10 ppm neomycin. The flasks were incubated for 48 hours at 32°C with constant rotational mixing at 180 rpm. Cultures were harvested by centrifugation at 14500 rpm for 20 min in conical tubes. The culture supernatants were used for assays. The cultivation media was an enriched semi-defined media based on MOPs buffer, with urea as major nitrogen source, glucose as the main carbon source, and supplemented with 1% soytone for robust cell growth. The nucleotide prop-mature sequence of the BspE04637 gene in plasmid pBN-BspE04637 is depicted in SEQ ID NO:7: CAATCACAGGCAGAAAAACAGGAGTACCTTGT  
 TCAATTTAATGACAAGGTAAATAAGGGGATATTAACGCATTTGGTGTAGATAACA  
 GTGATGTTCTTCATACTTACAATCTACTACCTGTTAATCTAGTAAAAATGACAGAGC  
 AGCAAGCGAAAGCATTACAAAATAATCCGCATATTAAGCGGTGGAACCTAACTTT  
 GAGGCGCAAGCATTTCGCTCAAACAGTACCGTGGGGAGTTCCTCATGTTCAAGGTA  
 CTGATGCTCATGCAGCGGGGCATACTGGAAGTGGTGTAAAAGTAGCTATACTCGATAC  
 GGGAATTGATCGAAATCATGAAGATTTAAATGTAAGAGGAGGGCACTCTGTATTCA  
 CAGACAGTGCAAACAGGGACCCTTATTACGATGGTAGTGGTCATGGCACGCACGTT  
 GCTGGAACAGTAGCGGCCTTAAATAATAGTGTAGGTGTGTTAGGTGTAGCCTACAAC  
 GCGGAGCTATATGCGGTAAAAGTATTAATAACAGTGGAAAGTGGTTCTTATGCGGG

AATCGCTGAAGGAATTGAATGGGCAGTACAAAACAATATGGACATTATAAATATGA  
 GTTTAGGAGGTTTCGATGAGTTCTTCTATTTTAGAAGAGTGGTGTAAATTGCCTACA  
 ATTCTGGTGTGCTTGTAGTGGCGGCAGCTGGAAATAGTGGACGTACGAATGGCCGA  
 GGGGATACTGTTGGTTATCCAGCAAATATGATTCTGTTATCGCAGTAGCAGCAGTT  
 GATTCTAGCAACAATCGTGCAAGTTTCTCTAGTACAGGACCTGCAGTGGAAATTGCT  
 GCACCAGGTGTAAACATTCTTAGTACGACTCCTGGAAACAGCTATGCATCTTACAAC  
 GGAACTTCGATGGCATCTCCACATGTAGCTGGTGTGGCAGCGCTTGTCTTGGCAGCG  
 AATCCGAACCTTTCAAACGTGGAGCTTCGTAATCGATTAAACGATACAGCGCAAAA  
 CTTAGGAGATGCAAACCACTTTGGAAACGGCCTAGTGCCTGCAGTTGATGCCATTAA  
 CGGCACTAGCTCCGGTGATAACGGTGGCGGAAGCGAACCAACAAAACCAGGTAACG  
 GCAAAGGAAACGGAAGAAAC.

[00187] The amino acid sequence of the BspE04637 precursor protein expressed from plasmid pBN-BspE04637 is depicted in SEQ ID NO:8 (the predicted pro-peptide is shown in underlined text): QSQAEEKQEYLVQFNDKVNKGILNAFGVDNSDVLHTYNLLPVNLVKMT  
EQQAKALQNNPHIKAVEPNFEAQAFAQTVPWGVPHVQGTDAHAAGHTGSGVKVAILD  
 TGIDRNHEDLNVRGGHVSFTDSANRDPYYDGS GHGTHVAGTVAALNNSVGLGVAYN  
 AELYAVKVLNNSGSGSYAGIAEGIEWAVQNNMDIINMSLGGSMSSSILEEWCNIAYNSG  
 VLVVAAAGNSGRTNNGRGTVDGYPYAKYDSVIAVAAVDSSNNRASFSSTGPAVEIAAPGV  
 NILSTTPGNSYASYNGTSMASPHVAGVAALVLAANPNLSNVELRNRLNDTAQNLGDAN  
 HFGNGLVRAVDANGTSSGDNGGGSEPTKPGNGKGNRN.

[00188] N and C-terminal analysis as described in Example 1 revealed that the isolated BspE04637 protein expressed from plasmid pBN-BspE04637 is a truncated version (280 amino acids), with the theoretical average mass of 28.5 kDa). The sequence of the truncated form of BspE04637 protein is shown in SEQ ID NO:9: AQTVPWGVPHVQGTDAHAAGHTGSGVKV  
 AILD TGIDRNHEDLNVRGGHVSFTDSANRDPYYDGS GHGTHVAGTVAALNNSVGLGV  
 AYN AELYAVKVLNNSGSGSYAGIAEGIEWAVQNNMDIINMSLGGSMSSSILEEWCNIA  
 YNSG VLVVAAAGNSGRTNNGRGTVDGYPYAKYDSVIAVAAVDSSNNRASFSSTGPAVEIAA  
 PGNILSTTPGNSYASYNGTSMASPHVAGVAALVLAANPNLSNVELRNRLNDTAQNLG  
 DAN HFGNGLVRAVDANGT.

## EXAMPLE 5

### Protease activity of BspE04637

[00189] The protease activity of BspE04637 protease (SEQ ID NO:9) was tested by measuring the hydrolysis of dimethyl casein (DMC) substrate. The reagent solutions used for the DMC assay were: 2.5% DMC (Sigma) in 100 mM Sodium Carbonate pH 9.5, 0.075% TNBSA (2,4,6-trinitrobenzene sulfonic acid, Thermo Scientific) in Reagent A. Reagent A: 45.4 g  $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$  (Merck) in 15 mL 4N NaOH to reach a final volume of 1000 mL in MQ water, Dilution Solution: 10 mM NaCl, 0.1 mM  $\text{CaCl}_2$ , and 0.005% Tween-80. Protease supernatants were diluted in Dilution Solution to appropriate concentration for the assay. A 96-well microtiter plate (MTP) was filled with 95  $\mu\text{l}$  DMC substrate followed by the addition of 5  $\mu\text{l}$  diluted protease supernatant. 100  $\mu\text{L}$  TNBSA in Reagent A was then added with slow mixing. Activity was measured at 405 nm over 5 min using a SpectraMax plate reader in kinetic mode at RT. The absorbance of a blank containing no protease was subtracted from the values. The activity was expressed as mOD/min. The protease activity curve for BspE04637 is shown in Figure 2. Using the DMC assay, the specific activity of BspE04637 protease was found to be 32 mOD/min/ppm. The specific activities of GG36 and BPN' proteases were found to be 54 and 23 mOD/min/ppm respectively under the same assay conditions.

### EXAMPLE 6

#### pH profile of BspE04637 protease

[00190] The pH dependence of proteolytic activity of BspE04637 protease (SEQ ID NO:9) was studied using azo-casein as substrate in a 50mM Acetate/Bis-Tris/HEPES/CHES buffer including 50 mM  $\text{CaCl}_2$ . The activity was measured at pH between 4 to 12 with 1 pH unit increments. One Protaxyme AK tablet (Megazyme, Ireland) was added to a glass test tube together with 1.9 mL of appropriate buffer and a magnet, followed by gentle hydration at 40°C for 5 min in a temperature controlled water bath fitted with magnetic stirrer. A 100  $\mu\text{L}$  sample of freshly prepared protease (diluted in deionised water to appropriate concentration for the assay) was added to the prehydrated substrate and reaction was carried out at 40°C for 10 min. To terminate the reaction, 10 mL of a 2 % w/v Tris buffer, pH 12 was added, solution was mixed, and the sample was immediately filtered through a Whatman No. 1 filter. The supernatant was collected, and the absorbance at 590 nm of the supernatant was measured to quantify the product of the reaction. The absorbance from a buffer-only control was subtracted, and the resulting values were converted to percentages of relative activity, by defining the activity at the optimal pH as 100%. BspE04637 was determined to maintain  $\geq 50\%$  of activity over the pH range of 6-

12, under the conditions of this assay.

### EXAMPLE 7

#### Temperature profile of BspE04637 protease

[00191] The temperature dependence of proteolytic activity of BspE04637 protease (SEQ ID NO:9) was studied using azo-casein as substrate in a 50mM Acetate/Bis-Tris/HEPES/CHES buffer including 50 mM CaCl<sub>2</sub> at pH 10. The activity was measured at temperatures between 30°C and 80°C with 10°C increments. One Protaxyme AK tablet (Megazyme, Ireland) was added to a glass test tube together with 1.9 mL of appropriate buffer and a magnet stirrer, followed by gentle hydration at set temperatures for 5 min in a temperature controlled water bath fitted with magnetic stirrer. A 100µl sample of freshly prepared protease (diluted in deionised water to appropriate concentration for the assay) was added to the prehydrated substrate and reaction was carried out at temperatures between 30°C and 80°C for 10 min. To terminate the reaction, 10 mL of a 2 % w/v Tris buffer pH 12 was added and solution was mixed and filtered immediately through a Whatman No. 1 filter. The supernatant was collected and the absorbance at 590 nm of the supernatant was measured to quantify the product of the reaction. The absorbance from a buffer-only control was subtracted from each sample reading, and the resulting values were converted to percentages of relative activity, by defining the activity at the optimal temperature at 100%. BspE04637 was determined to retain ≥50% activity over a range of 60-80°C, under the conditions of this assay.

### EXAMPLE 8

#### Stability evaluation of proteases

[00192] The stability of BspE04637 (SEQ ID NO:9) and FNA (SEQ ID NO:10) proteases was determined under various conditions.

[00193] SEQ ID NO:10 sets forth the sequence of FNA protease: AQSVPYGVVSQIKAPALH SQGYTGSNVKVAVIDSGIDSSHPDLKVAGGASMVPSETNPFQDNNSHGTHVAGTVAAL NNSIGVLGVAPSASLYAVKVLGADGSGQYSWIINGIEWAIANNMDVINMSLGGPSGSAA LKAAVDKAVASGVVVVAAAGNEGTS GSSSTVGYPGKYPSVIAVGAVDSSNQRASFSSV GPELDVMAPGVSIQSTLPGNKYGALNGTSMASPHVAGAAALILSKHPNWTNTQVRSSL ENT TTKLGDSFY YGKGLINVQAAAQ.

[00194] Stability was tested under four stress conditions shown below by measuring the residual proteolytic activity following incubation at set temperatures.

LAS/EDTA: 0.02% LAS, 2.1mM EDTA in 50mM HEPES pH8, 0,005% Tween 80

OMO HDL: 10% OMO Klein & Krachtig (proteases present in the detergent was inactivated prior to use)

[00195] For stressed conditions, diluted enzyme sample was mixed in stress buffers/detergent in a 96-well PCR plate and incubated at 30°C, 40°C, 50°C, 60°C and 75°C for 20 min using a Tetrad2 Thermocycler. For the unstressed condition, enzyme was assayed immediately after mixing with stress media to establish a baseline (initial activity). Protease activity under stressed and unstressed conditions was measured by either the hydrolysis of AAPF-pNA or DMC substrate assays described previously. Percent residual activities were calculated by taking a ratio of the stressed to unstressed activity at each temperature and multiplying it by 100. The percent remaining activity for each protease is shown on Tables 1-2 for each condition run at the various temperatures.

Table 1: Stability of proteases in LAS/EDTA						
Enzyme	Unstressed	Incubation temperature				
		30°C	40°C	50°C	60°C	75°C
BspE04637	100	>90	>90	67	21	15
FNA	100	>90	84	7	4	1

Table 2: Stability of proteases in 10% OMO HDL						
Enzyme	Unstressed	Incubation temperature				
		30°C	40°C	50°C	60°C	75°C
BspE04637	100	>90	>90	>90	>90	0
FNA	100	>90	91	71	1	1

## EXAMPLE 9

### Comparison of BspE04637 protease to Related Molecules

#### Identification of Homologous Proteases

[00196] Related proteins were identified by a BLAST search (Altschul et al., Nucleic Acids Res, 25:3389-402, 1997) using the mature protein amino acid sequences for BspE04637 protease (SEQ ID NO:9) and SWT183\_1430046 protease (SEQ ID NO: 6) against the NCBI non-redundant protein database and a subset are shown on Tables 3A and 4A, respectively. A similar search was run against the Genome Quest Patent database with search parameters set to default values using the mature protein amino acid sequences for BspE04637 protease (SEQ ID NO:9) ) and SWT183\_1430046 protease (SEQ ID NO: 6) as the query sequence, and a subset are

shown in Tables 3B and 4B, respectively. Percent identity (PID) for both search sets is defined as the number of identical residues divided by the number of aligned residues in the pairwise alignment. Value labeled “Sequence length” on tables corresponds to the length (in amino acids) for the proteins referenced with the listed Accession numbers, while “Aligned length” refers to sequence used for alignment and PID calculation.

Accession #	PID	Organism	Sequence Length	Alignment Length
BAD02409.1	99	<i>Bacillus sp. KSM-LD1</i>	404	307
WP_026690432	77	<i>Bacillus aurantiacus</i>	377	278
WP_035661169	76	<i>Bacillus akibai JCM 9157</i>	378	280
ADC50469.1	75	<i>Bacillus pseudofirmus OF4</i>	379	279
WP_047973355	75	<i>Bacillus sp. LL01</i>	377	279
ERN55058/ WP_022626565.1	75	<i>Bacillus marmarensis DSM 21297</i>	291	279
WP_026475840	74	<i>Alkaliphilus transvaalensis</i>	382	280
WP_027963976	70	<i>Halalkalibacillus halophilus</i>	372	277
WP_026691136	69	<i>Bacillus aurantiacus</i>	376	276
WP_017185212.1	68	<i>Alkalibacillus haloalkaliphilus</i>	374	276
WP_010192403.1	64	<i>Bacillus sp. m3-13</i>	381	277
AAC43580.1	64	<i>Bacillus sp. SprC</i>	378	277
AAC43581.1	64	<i>Bacillus sp. SprD</i>	379	277
BAD11988.2	63	<i>Bacillus sp. KSM-LD1 SA</i>	376	277
BAD21128.1	63	<i>Bacillus sp. KSM-LD1 SB</i>	377	277
CAJ70731.1	59	<i>Bacillus licheniformis</i>	379	271
CAA24990.1	59	<i>Bacillus amyloliquefaciens</i>	376	271
P29600	58	<i>Bacillus Lentus Subtilisin</i>	269	273
ADC49870.1	58	<i>Bacillus pseudofirmus OF4</i>	374	273
WP_015383267.1	57	<i>Bacillus subtilis</i>	381	271
BAA06157.1	56	<i>Bacillus sp. Sendai</i>	382	272

Patent ID #	PID	Organism	Sequence Length	Alignment Length
JP2003325186-0001	99.3	<i>Bacillus sp.; strain KSM-LD1</i>	404	307
WO2015038792-0019	66.8	<i>Bacillus horikoshii DSM 6951</i>	275	277
WO2015038792-0010	66.4	<i>Bacillus horikoshii DSM 8719</i>	275	277
WO2015038792-0016	66.1	<i>Bacillus horikoshii DSM 9712</i>	275	277

Table 3B: List of sequences with percent identity to BspE04637 protein identified from the Genome Quest database				
Patent ID #	PID	Organism	Sequence Length	Alignment Length
WO2015038792-0013	65.3	<i>Bacillus horikoshii</i> DSM 9711	275	277
WO2015089447-0085	65.0	<i>Bacillus</i> sp. m3-13	275	277
JP2012524542-0052	64.8	<i>Alkaliphilus transvaalensis</i>	274	278
WO9406915	64.6	<i>Bacillus</i> sp	275	277
JP2004154003-0007	63.9	<i>Bacillus</i> sp. KSM-LD1	276	277
JP2004313043-0022	62.8	<i>Bacillus</i> sp. KSM-9865	275	277
JP2004313043-0021	62.1	<i>Bacillus</i> sp. KSM-KP43	275	277
JP2013500714-0395	59.4	<i>Bacillus subtilis</i>	275	271

Table 4A: List of sequences with percent identity to SWT183\_1430046 protein identified from the NCBI non-redundant protein database

Accession #	PID	Organism	Sequence Length	Alignment Length
BAD02409	97.1	<i>Bacillus</i> sp. KSM-LD1	404	307
WP_026690432	77.0	<i>Bacillus aurantiacus</i>	377	278
WP_035661169	75.4	<i>Bacillus akibai</i> JCM 9157	378	280
WP_047973355	74.9	<i>Bacillus</i> sp. LL01	377	279
WP_012957833	74.2	<i>Bacillus pseudofirmus</i> OF4	379	279
ERN55058/ WP_022626565.1	73.8	<i>Bacillus marmarensis</i> DSM 21297	291	279
WP_026475840	73.6	<i>Alkaliphilus transvaalensis</i>	382	280
WP_027963976	70.0	<i>Halalkalibacillus halophilus</i>	372	277
WP_026691136	69.2	<i>Bacillus aurantiacus</i>	376	276
WP_017185212	68.8	<i>Alkalibacillus haloalkaliphilus</i>	374	276
WP_027965007	68.5	<i>Halalkalibacillus halophilus</i>	377	279
WP_047970940	67.1	<i>Bacillus</i> sp. LL01	377	277
WP_048311846	66.2	<i>Anaerobacillus macyae</i>	382	275
WP_048310299	66.2	<i>Anaerobacillus macyae</i>	379	275
AAC43580	65.3	<i>Bacillus</i> sp. SprC	378	277
AAC43581	65.0	<i>Bacillus</i> sp. SprD	379	277
WP_010192405	65.0	<i>Bacillus</i> sp. m3-13	379	277
WP_026475841	64.4	<i>Alkaliphilus transvaalensis</i>	376	278
BAD11988	64.3	<i>Bacillus</i> sp. KSM-LD1	376	277
BAD21128	63.9	<i>Bacillus</i> sp. KSM-LD1	377	277
WP_034327386	63.6	<i>Alkaliphilus transvaalensis</i>	382	280
AHZ65108	62.9	<i>Bacillus subtilis</i>	297	237
WP_029281843	62.7	<i>Bacillus indicus</i> LMG 22858	375	276
WP_029565418	62.3	<i>Bacillus cibi</i>	375	276
WP_044803561	61.6	<i>Bacillus amyloliquefaciens</i>	333	232
ABD33463	60.7	<i>Bacillus</i> sp. hr08	355	257



Table 4A: List of sequences with percent identity to SWT183_1430046 protein identified from the NCBI non-redundant protein database				
WP_048353666	60.7	<i>Bacillus sp. GO-13</i>	378	272
WP_003327717	60.5	<i>Bacillus atrophaeus UCMB-5137</i>	382	271
ADI24411	60.3	<i>Bacillus subtilis</i>	347	242
WP_046128901	60.3	<i>Bacillus sp. TH008</i>	378	272
WP_024712916	60.1	<i>Bacillus tequilensis</i>	381	271
WP_036843491	60.0	<i>Pontibacillus marinus</i>	363	275

Table 4B: List of sequences with percent identity to SWT183_1430046 protein identified from the Genome Quest database				
Patent ID #	PID	Organism	Sequence Length	Alignment Length
JP2003325186-0001	97.1	<i>Bacillus sp.KSM-LD1(1)</i>	404	307
WO2015038792-0019	67.2	<i>Bacillus horikoshii DSM 6951</i>	275	277
WO2015038792-0016	66.8	<i>Bacillus horikoshii DSM 9712</i>	275	277
WO2015038792-0010	66.4	<i>Bacillus horikoshii DSM 8719</i>	275	277
WO2015089447-0085	66.1	<i>Bacillus sp. m3-13</i>	275	277
WO2015038792-0013	65.7	<i>Bacillus horikoshii DSM 9711</i>	275	277
WO2015038792	65.6	<i>Bacillus sp.LG12 Synthetic</i>	275	276
WO2015038792-0004	65.3	<i>Bacillus sp.LG12</i>	381	277

#### Alignment of Homologous Sequences

[00197] An alignment of the amino acid sequences of the mature BspE04637 protease (SEQ ID NO:9) and mature SWT183\_1430046 protease (SEQ ID NO:6) with the sequences of the mature forms of select subtilisins from Tables 3A and 4A is shown in Figure 3. The sequences were aligned with default parameters using the MUSCLE program from Geneious software (Biomatters Ltd.) (Robert C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucl. Acids Res. (2004) 32 (5): 1792-1797).

[00198] A phylogenetic tree for amino acid sequences of the mature forms of the subtilisins from Figure 5 was built using the Geneious Tree builder program, and is set forth in Figure 3.

### EXAMPLE 10

#### Unique Features of the BspE04637-clade of Subtilisins

[00199] The Figure 3 alignment was reviewed for unique sequence similarities across the BspE04637-clade of subtilisins. The BspE04637-clade of subtilisins is characterized by a common motif over the sequence that begins with the catalytic Aspartic acid (D32) of the serine protease triad and ends with the catalytic Histidine (H66) of the catalytic triad according to BspE04637 numbering. When compared to the amino acid sequence of BPN', the motif includes

one insertion of 2 amino acid residues. The motif can be characterized by Motif 1: DTGIXXXH XDLXXXGGXSVFXXXXXXXXXXDXXGH (SEQ ID NO:31), wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid; Motif 2: DTGIXXXHXDLXXXGGXSVFXXXXXXXXDPXXDXXGH (SEQ ID NO:32), wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid; Motif 3: DTGIXXXHXDLNVXGGXSVFXXXXXXXXXX DXXGH (SEQ ID NO:33), wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid; Motif 4: DTGIXXXHXDLNVXGGXSVFXX XXXXDPXX DXXGH (SEQ ID NO:34), wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid; Motif 5: DTGIDXXHXDLNVX GGXSVFXX XXXXXXXXXXXDXXGH (SEQ ID NO:35), wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid; Motif 6: DTGIDXNHX DLNVRGGXSVFTXXXXXXXXDPXXDXXGH (SEQ ID NO:36), wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid; or Motif 7: DTGIDXNHXDLNVRGGXSVFTXXXXXXXXDPYYDXXGH (SEQ ID NO:37), wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid. Figure 3 includes a box around the motif.

**[00200]** The insertion is expected to occur in a loop common to the overall tertiary fold of subtilisin enzymes. The location of the insertion is illustrated in Figure 5 using the known structure of subtilisin BPN' as a reference. The motif segment is highlighted in black using the BPN' subtilisin structure as a reference (in light gray). The catalytic Aspartic acid and Histidine residue side chains, of the catalytic triad common to all serine proteinases are shown as sticks. The segment that connects the catalytic Aspartic Acid (D32) and Histidine (H66) (based on BspE04637 numbering) of the subtilisin catalytic triad comprises two loops and the outermost strand of the central beta sheet. This strand includes the GGXS portion of the BspE04637-clade motif and is indicated by an arrow in Figure 5. The insertion is proposed to occur in the loop indicated by the arrow, wherein the insertion enlarged the second loop. The loop containing the insertion follows the GGXS strand and leads into the catalytic Histidine.

**[00201]** The BspE04637, SWT183\_1430046, and BAD02409 subtilisins which have been identified as BspE04637-clade of subtilisins based on the shared sequence motif set forth above, also cluster together in the phylogenetic tree that was built using various bacterial subtilisins and

which is set forth in Figure 4.

**[00202]** While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein can be employed in practicing the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

## CLAIMS

We claim:

1. A BspE04637-clade of subtilisins comprising a subtilisin or recombinant polypeptide or active fragment thereof comprising one or more motif selected from a:
  - i. DTGIXXXHXDLXXXGGXSVFXXXXXXXXXXDXXGH (SEQ ID NO:31) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid;
  - ii. DTGIXXXHXDLXXXGGXSVFXXXXXXXXDPXXDXXGH (SEQ ID NO:32) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid;
  - iii. DTGIXXXHXDLNVXGGXSVFXXXXXXXXXXDXXGH (SEQ ID NO:33) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid;
  - iv. DTGIXXXHXDLNVXGGXSVFXXXXXXXXDPXXDXXGH (SEQ ID NO:34) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid;
  - v. DTGIDXXHXDLNVXGGXS VFXXXXXXXXXXDXXGH (SEQ ID NO:35) motif, wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid;
  - vi. DTGIDXNHXDLNVRGGXSVFTXXXXX DPXXDXXGH (SEQ ID NO:36) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid; and
  - vii. DTGIDXNHXDLNVRGGXSVFTXXXXX DPYYDXXGH (SEQ ID NO:37) motif, wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid.
2. The BspE04637-clade of subtilisins according to Claim 1, wherein the subtilisin or recombinant polypeptide or active fragment thereof further comprises an amino acid sequence having at least 70% amino acid sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:3, 6 and 9.
3. The BspE04637-clade of subtilisins according to Claim 1 or 2, with the proviso that the subtilisin and/or recombinant polypeptide or active fragment thereof does not comprise

BAD02409 or JP2003325186-0001.

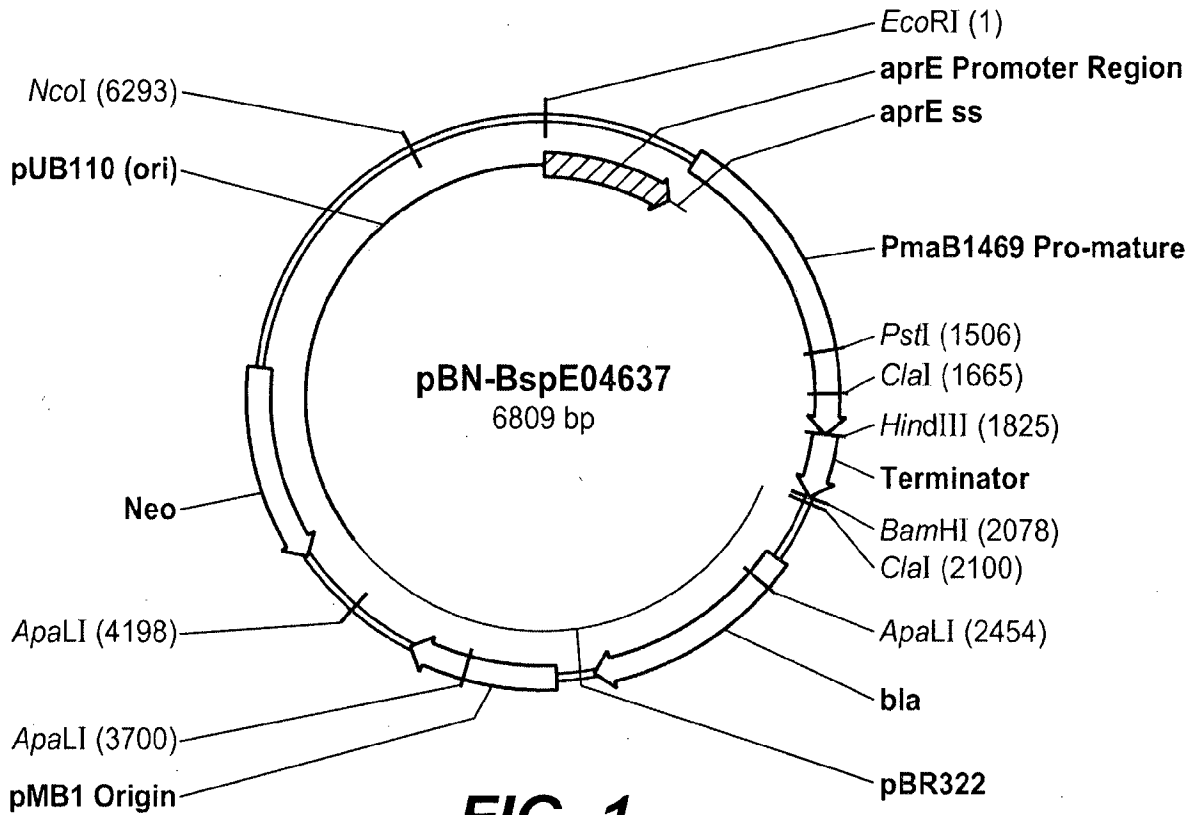
4. The BspE04637-clade of subtilisins of any of the above Claims, wherein the subtilisin and/or recombinant polypeptide or active fragment has protease activity.
5. A recombinant polypeptide or an active fragment thereof, comprising an amino acid sequence having at least 80% amino acid sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:3, 6 and 9.
6. A recombinant polypeptide or active fragment thereof of the BspE04637-clade.
7. The recombinant polypeptide or active fragment thereof of Claim 6, wherein the recombinant polypeptide or active fragment comprises one or more motif selected from a:
  - i. DTGIXXXHXDLXXXGGXSVFXXXXXXXXXXDXXGH (SEQ ID NO:31) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid;
  - ii. DTGIXXXHXDLXXXGGXSVFXXXXXXXXDPXXDXXGH (SEQ ID NO:32) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid;
  - iii. DTGIXXXHXDLNVXGGXSVFXXXXXXXXXXDXXGH (SEQ ID NO:33) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid;
  - iv. DTGIXXXHXDLNVXGGXSVFXXXXXXXXDPXXDXXGH (SEQ ID NO:34) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid;
  - v. DTGIDXXHXDLNVXGGXSVFXXXXXXXXXXDXXGH (SEQ ID NO:35) motif, wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid;
  - vi. DTGIDXNHXDLNVRGGXSVFTXXXXXXXXDPXXDXXGH (SEQ ID NO:36) motif, wherein the initial D is the active site Aspartic acid, the terminal H is the active site Histidine, and X is any amino acid; and
  - vii. DTGIDXNHXDLNVRGGXSVFTXXXXXXXXDPYYDXXGH (SEQ ID NO:37) motif, wherein the initial D is the active site Aspartic acid and the terminal H is the active site Histidine, and X is any amino acid.
8. The recombinant polypeptide or active fragment thereof of Claim 6 or 7, wherein the

- recombinant polypeptide or active fragment thereof further comprises an amino acid sequence having at least 70% amino acid sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:3, 6 and 9.
9. The recombinant polypeptide or active fragment thereof of any one of Claims 5-8, with the proviso that the subtilisin and/or recombinant polypeptide or active fragment thereof does not comprise BAD02409 or JP2003325186-0001.
  10. The recombinant polypeptide or active fragment thereof of any one of Claims 5-9, wherein the polypeptide has protease activity.
  11. The recombinant polypeptide or active fragment thereof of Claim 10, wherein the protease activity comprises casein hydrolysis activity or dimethylcasein hydrolysis activity.
  12. The recombinant polypeptide or active fragment thereof of any one of Claims 5-11, wherein the polypeptide retains at least 50% of its maximal protease activity at a pH range of 6 to 12 and/or a temperature range of 55°C to 80°C.
  13. The recombinant polypeptide or active fragment thereof of any one of Claims 5-12, wherein the polypeptide retains at least 60% activity after 20 minutes at 50°C under stressed conditions.
  14. The recombinant polypeptide or active fragment thereof of claim 13, wherein the stressed condition is in an LAS/EDTA or OMO HDL assay.
  15. The recombinant polypeptide or active fragment thereof of any one of Claims 5-14, wherein the polypeptide has cleaning activity in a detergent composition.
  16. The recombinant polypeptide or active fragment thereof of Claim 15, wherein the detergent composition is an automatic dish washing detergent and/or the cleaning activity comprises hydrolysis of an egg yolk substrate.
  17. The recombinant polypeptide or active fragment thereof of Claim 15, wherein the detergent composition is a laundry detergent and/or the cleaning activity comprises hydrolysis of a substrate selected from the group consisting of blood, milk, ink and combinations thereof.
  18. The recombinant polypeptide or active fragment thereof of Claim 17, wherein the laundry detergent is a liquid laundry detergent or a powder laundry detergent.
  19. A composition comprising a surfactant and the subtilisin or recombinant polypeptide or active fragment thereof of any one of claims 1-4 or the recombinant polypeptide or active fragment thereof of any one of Claims 5-18.

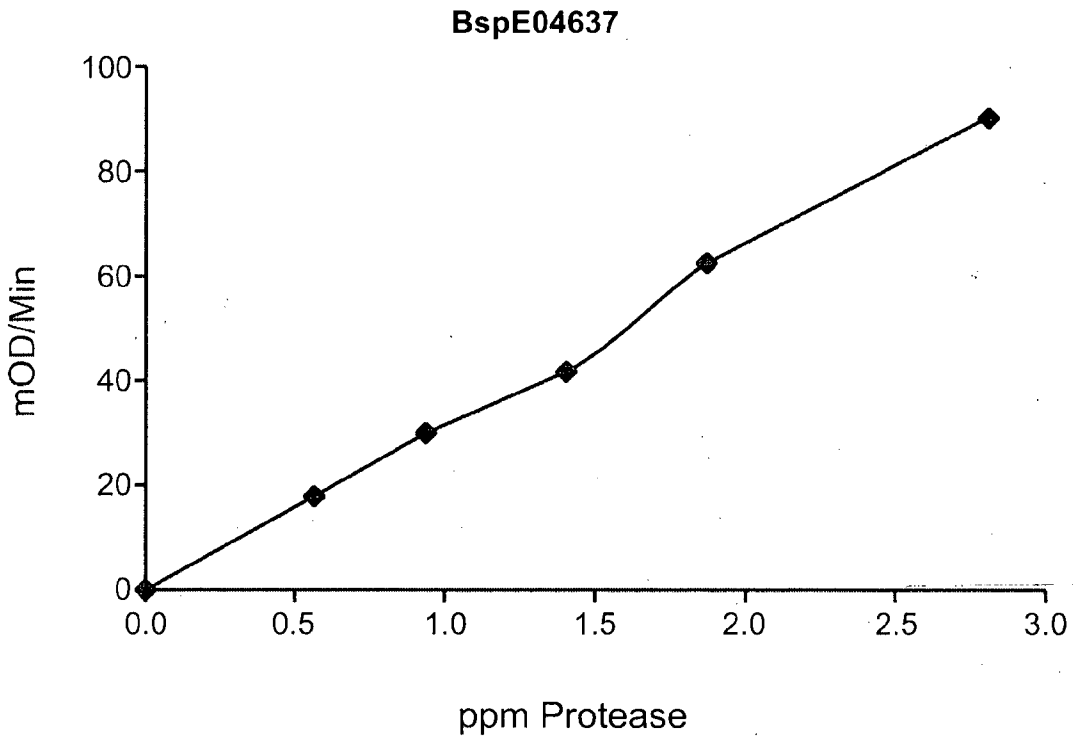
20. The composition of Claim 19, wherein the surfactant is selected from the group consisting of an anionic surfactant, a cationic surfactant, a zwitterionic surfactant, an ampholytic surfactant, a semi-polar non-ionic surfactant, and a combination thereof.
21. The composition of Claim 19 or 20, wherein the composition is a detergent composition.
22. The composition of Claim 21, wherein the detergent composition is selected from the group consisting of a laundry detergent, a fabric softening detergent, a dishwashing detergent, and a hard-surface cleaning detergent.
23. The composition of any one of claims 19-22, wherein said composition further comprises at least one calcium ion and/or zinc ion; at least one stabilizer; from about 0.001% to about 1.0 weight % of said recombinant polypeptide; at least one bleaching agent; at least one adjunct ingredient; and/or one or more additional enzymes or enzyme derivatives selected from the group consisting of acyl transferases, alpha-amylases, beta-amylases, alpha-galactosidases, arabinosidases, aryl esterases, beta-galactosidases, carrageenases, catalases, cellobiohydrolases, cellulases, chondroitinases, cutinases, endo-beta-1, 4-glucanases, endo-beta-mannanases, esterases, exo-mannanases, galactanases, glucoamylases, hemicellulases, hyaluronidases, keratinases, laccases, lactases, ligninases, lipases, lipoxygenases, mannanases, oxidases, pectate lyases, pectin acetyl esterases, pectinases, pentosanases, peroxidases, phenoloxidases, phosphatases, phospholipases, phytases, polygalacturonases, proteases, pullulanases, reductases, rhamnogalacturonases, beta-glucanases, tannases, transglutaminases, xylan acetyl-esterases, xylanases, xyloglucanases, xylosidases, metalloproteases, additional serine proteases, and combinations thereof.
24. The composition of any of Claims 19-23, wherein said composition contains phosphate or is phosphate-free and/or contains borate or is borate-free.
25. The composition of any of Claims 19-24, wherein said composition is a granular, powder, solid, bar, liquid, tablet, gel, paste or unit dose composition.
26. The composition of any one of Claims 19-25, wherein said composition is formulated at a pH of from about 8 to about 12.
27. A method of cleaning, comprising contacting a surface or an item in need of cleaning with the subtilisin or recombinant polypeptide or active fragment thereof of any one of claims 1-4; the recombinant polypeptide or active fragment thereof of any one of Claims 5-18; or the composition of any one of Claims 19-26; and optionally further comprising the step of

- rinsing said surface or item after contacting said surface or item with said recombinant polypeptide or composition.
28. The method of Claim 27, wherein said item is dishware or fabric.
  29. A polynucleotide comprising a nucleic acid sequence that encodes the subtilisin or recombinant polypeptide or active fragment thereof of any one of claims 1-4; or the recombinant polypeptide or active fragment thereof of any one of Claims 5-18.
  30. The polynucleotide of Claim 29, wherein said polynucleotide comprises a nucleic acid sequence having at least 70% nucleic acid identity to SEQ ID NO:1, 4 or 7.
  31. An expression vector comprising the polynucleotide of Claim 29 or 30.
  32. A host cell comprising the vector of Claim 31.
  33. The host cell of Claim 32, wherein the host cell is of a species selected from *Bacillus spp.*, *Streptomyces spp.*, *Escherichia spp.*, *Aspergillus spp.*, *Trichoderma spp.*, *Pseudomonas spp.*, *Corynebacterium spp.*, *Saccharomyces spp.*, and *Pichia spp.*
  34. The host cell of Claim 33, wherein said *Bacillus spp.* is *Bacillus subtilis*.
  35. A method for producing the subtilisin or recombinant polypeptide or active fragment thereof of any one of Claims 1-4; or the recombinant polypeptide or active fragment thereof of any one of Claims 5-18 comprising:
    - (a) stably transforming the host cell of any one of Claims 32-34 with the expression vector of Claim 31;
    - (b) cultivating said transformed host cell under conditions suitable for said host cell to produce said subtilisin or polypeptide; and
    - (c) recovering said subtilisin or polypeptide.
  36. The method of Claim 35, wherein said expression vector comprises a heterologous polynucleotide sequence encoding a heterologous pro-peptide.
  37. The method of Claim 35 or 36, wherein said expression vector comprises one or both of a heterologous promoter and a polynucleotide sequence encoding a heterologous signal peptide.
  38. A composition comprising the subtilisin or recombinant polypeptide or active fragment thereof of any one of Claims 1-4; or the recombinant polypeptide or active fragment thereof of any one of Claims 5-18, wherein said composition is an animal feed, contact lens cleaning, wound cleaning, or textile, leather or feather processing composition.





**FIG. 1**



**FIG. 2**

1	AQTVPWGVPHVQGTDAHAAGHTGSGVKVAILLDTGIDRNHEDL--NVRGGHSVFTDSANRD	BspE04637
	AQTVPWGVPHVQGTDAHAAGHTGSGVKVAILLDTGIDRNHEDL--NVRGGHSVFTDSANSD	SWT183_1430046
	AQTVPWGVPHVQGTDAHAAGHTGSGVKVAILLDTGIDRNHEDL--NVRGGHSVFTDSANRD	B_sp_BAD02409.1
	GQTVPWGIPHVQGTAAQDAGYTGAGLKVAILDTGIDRNHEDLNFENVKGGHSVFTDAANRD	B_aurantiacus_WP_026690432.1
	AESVPWGVPHVQGTAAQANGFTGSGVKVAILLDTGIDLSHEDLSANVKGFSVFDAAANRD	B_akibai_WP_035661169.1
	AQNVPGIPRVQGTDAQNAQYTGNGVKVAILLDTGIDRSHPDLTANVQGGHSVFTDSANRD	B_pseudofirmus_OF4_ADC50469
	AQTTVPWGIIPRVQGTAAQDAGFTGNGVKVAILLDTGIDRNHPDLSANVKGGHSVFTDSANSD	B_spLL01_WP_047973355.1
	AQNVPGIPRVQGTDAQNAQYTGNGVKVAILLDTGIDRSHPDLTANVQGGHSVFTDSANRD	B_marmarensis_WP_022626565_ERN55058.1
	SQTVPWGISHVQAITAHNSGYTGENVKVAILLDTGIDNKHEDLFSNVKGGYSVFTDSANSD	A_transvaalensis_WP_026475840.1
	AQQTVPWGIIPRVEGTTSQNNGYTDGIDVAVLDTGIDRSHVLD--NVSGGYSVFGDS----	H_halophilus_WP_027963976.1
	AQSVPWGISRVQAPAAHNRGLTSGVKVAVLDTGI-STHPDL--NIRGGASFVPEPS---	B_lentus_P29600
	NQVTPWGITRVQAPTAWTRGYTGTGVRVAVLDTGI-STHPDL--NIRGGVSFVPEPS---	B_sp_Sendai_BAA06157
	AQTVPWGIPYIYSDVVHRQGYFGNGVKVAVLDTGV-APHPDL--HIRGGVSFISTENT---	B_pseudofirmus_ADC49870
	AQSVPYGVSQIKAPALHSQGYTGSNVKVAVIDSGIDSSHPDL--KVAGGASMPSETN---	B_amyloliquefaciens_CAA24990
	AQSVPYGISQIKAPALHSQGYTGSNVKVAVIDSGIDSSHPDL--NVRGGASFVPEPN---	B_subtilis_WP_015383267
	AQTVPYGIPLIKADKVAQGFKGANVKVAVLDTGIQASHPDL--NVVGGASFVAGEAY---	B_licheniformis_CAJ70731.1
	SQTVPYGVPHIKADVHASNQVTGNGVKVAILLDTGIDAHEDL--RVVGGASFVAGEPN---	B_sp_KSM-LD1_SB_BAD21128
	AQTTVPWGVTHINAHRAHSSGYTSGVKVAILLDTGIHASHPDL--NVRGGASFISGESN---	B_sp_KSM-LD1_SA_BAD11988
	AQTVPWGIPHIKADKAHASGYTSGVKVAVLDTGIDANHADL--NVKGGASFVSPEPN---	B_sp_m3_13_WP_010192403
	AQTVPWGIPHIKADKAHAAGYTGSGVKVAILLDTGIDANHADL--NVKGGASFVSPEPN---	Bacillus_sp_SprC_AAC43580

FIG. 3A

PYYDGSCHGTHVAGTVAALNNSVGLGVAYNAELIYAVKVLNNSGSGSYAGIAEGIEWAVQ  
 PYYDGSCHGTHVAGTVAALNNSVGLGVAYNAELIYAVKVLNNSGSGSYAGIAEGIEWAVQ  
 PYYDGSCHGTHVAGTVAALNNSVGLGVAYNAELIYAVKVLNNSGSGSYAGIAEGIEWAVN  
 PFYDADGCHGTHVAGTVAADVNDLGVVGVASQAELIYAVKVLNNSGSGSYAGIAEGIEWSIN  
 PYYDANGCHGTHVAGTVAAVNNLGVVGVAYQADLYAVKVLNNDGSGSYAGIARGIEWSVQ  
 PFFDGDGCHGTHVAGTVAAVNNDIGVVGVASEADLYAVKVLNNSGSGSYAGIAEGIEWSIN  
 PFFDGDGCHGTHVAGTVAAVNNDIGVIGVASEASLYAVKVLNNSGSGSYAGIAEGIEWAIN  
 PFFDGDGCHGTHVAGTVAAVNNDIGVVGVASEADLYAVKVLNNSGSGSYAGIAEGIEWSIN  
 PFYDGNHGCHGTHVAGTVAALDNNLGVVGVAYNADLYAVKVLNNSGSGSYDGLAQQIEWAVN  
 PYYDGDGCHGTHVAGTIAALDNTTGVVGVAPDANLYAVKVLNNSGSGSYDGLAQQIEWSII  
 -TQDCNGHGTHVAGTIAALNNSIGVVLGVAPSAELIYAVKVLGASGSGSVSSIAQGLEWAGN  
 -YQDCNGHGTHVAGTIAALNNSIGVVGVAPNAELIYAVKVLGANGSGSVSSIAQGLQWTAQ  
 -YVDYNGHGTHVAGTVAALNNSYGVVGVAPGAELIYAVKVLDRNGSGSHASIAQGLEWAMN  
 PFQDNNSHGTHVAGTVAALNNSIGVVLGVAPSAELIYAVKVLGADGSGQYSWIINGIEWAIA  
 PYQDSSHGTHVAGTIAALNNTTIGVVLGVAPNASLYAVKVLDDSTGSGQYSWIINGIEWAIS  
 -NTDGNHGCHGTHVAGTVAALDNTTGVVGVAPSVSLYAVKVLNNSGSGSYSGIVSGIEWATT  
 ALQDGNHGCHGTHVAGTVAALNNQVGLGVAYDLDLYAVKVLGADGSGTSLGIAQIEWSIA  
 PYIDSNHGCHGTHVAGTVAALNNTVGLGVAYNAELIYAVKVLASGSGTSLGIAQIEWSIA  
 ALQDGNHGCHGTHVAGTVAALNNTTGVVGVAYNADLYAVKVLASGSGTSLGIAQIEWSIA  
 ALQDGNHGCHGTHVAGTVAALNNTTGVVGVAYNADLYAVKVLASGSGTSLGIAQIEWSIS

BspE04637  
 SWT183\_1430046  
 B\_sp\_BAD02409.1  
 B\_aurantiacus\_WP\_026690432.1  
 B\_akibai\_WP\_035661169.1  
 B\_pseudofirmus\_OF4\_ADC50469  
 B\_spLL01\_WP\_047973355.1  
 B\_marmarensis\_WP\_022626565\_ERN55058.1  
 A\_transvaalensis\_WP\_026475840.1  
 H\_halophilus\_WP\_027963976.1  
 B\_lentus\_P29600  
 B\_sp\_Sendai\_BAA06157  
 B\_pseudofirmus\_ADC49870  
 B\_amyloliquefaciens\_CAA24990  
 B\_subtilis\_WP\_015383267  
 B\_licheniformis\_CAJ70731.1  
 B\_sp\_KSM-LD1\_SB\_BAD21128  
 B\_sp\_KSM-LD1\_SA\_BAD11988  
 B\_sp\_m3\_13\_WP\_010192403  
 Bacillus\_sp\_SprC\_AAC43580

FIG. 3B

BspE04637	NNMDIINMSLGGSMSSSILEEWCNIAYNISGVLVAAAAGNSGRTNCRGDTVGYPAKYDSVI
SWT183_1430046	NNMDIINMSLGGSMSSSILEEWCNIAYNISGVLVAAAAGNSGRTNCRGDTVGYPAKYDSVI
B_sp_BAD02409.1	NGMDIINMSLGGSMSSSILEEWCNIAYNISGVLVAAAAGNSGRTNCRGDTVGYPAKYDSVI
B_aurantiacus_WP_026690432.1	NGMDIINMSLGGSSSILKQFSDLAYEEGLVAAAAGNSGRRGGNDTVGYPAKYDSVI
B_akibai_WP_035661169.1	NGMDIVNMSLGGTSSSILKEWSDLAYAQVLLVAAAAGNSGTRPGRGDNVGYPAKYDSVI
B_pseudofirmus_OF4_ADC50469	NGMDIINMSLGGSQSSAILKEFSDLAYAEGLVAAAAGNSGRRGGNDTVGYPAKYESVI
B_spLL01_WP_047973355.1	NDIDIINMSLGGSQSSAILKEFSDLAYAEGLVAAAAGNSGTRSGRNDTVGYPAKYDSVI
B_marmarensis_WP_022626565_ERN55058.1	NGMDIINMSLGGSSSAILKEFSDLAYAEGLVAAAAGNSGRRGGNDTVGYPAKYESVI
A_transvaalensis_WP_026475840.1	NGMDIINMSLGGSSSAILLENMVAANNAGVLLIAAAGNSGNIIFGWGDTVGYPAKYGSVM
H_halophilus_WP_027963976.1	NGMDIINMSLGGPSSSILQQYSDLAYNNGILVAAAAGNSGNSWGWGDTVGYPAKYDSVI
B_lentus_P29600	NGMHVANLSLGGSPSATLEQAVNSATSRGVLVVAASGNSGA-----GSISYPARYANAM
B_sp_Sendai_BAA06157	NNITHVANLSLGGSPVGSQTLEAVNQATNAGVLVVAATGNNGS-----GTVSYPARYANAL
B_pseudofirmus_ADC49870	NGMDIANMSLGGSPGTTLQLAADRARNAGVLLIGAAGNSGQ-QGGSNMNGYPARYASVM
B_amycoliquefaciens_CAA24990	NNMDVINMSLGGPSSSAAALKAADVKAVASGVVVAAGNEGT-SGSSSTVGYPGKYPSVI
B_subtilis_WP_015383267	NNMDVINMSLGGPSSGSTALLKTVDKAVSSGIVVAAAAGNEGS-SGSTSTVGYPAKYPSVI
B_licheniformis_CAJ70731.1	NGMDVINMSLGGASGSTALLKQAVDNAYARGVVVAAAAGNSGS-SGNTNTIGYPAKYDSVI
B_sp_KSM-LD1_SB_BAD21128	NNMDVINMSLGGSTGSTALLKQAADNAYNSGLVAAAAGNSGDFFLINTIGYPARYDSVI
B_sp_KSM-LD1_SA_BAD11988	NKMDVINMSLGGSSGSTALLQRAVDNAYRNIVVAAAAGNSGA-QGNRNTIGYPARYSSVI
B_sp_m3_13_WP_010192403	NDMDVINMSLGGSTGSTALLQACDNAYASGIVVAAAAGNSGS-KGKRNTMGYPARYSSVI
Bacillus_sp_SprC_AAC43580	NGMNVINMSLGGSSGSTALLQACNNAYNRGIVVIAAAGNSGS-SGNRNTMGYPARYSSVI

**FIG. 3C**

BspE04637	AVAAVDSSNNRASFSSTGPAVEIAAPGVNILSTTPGNSYASVNGTSMASPHVAGVAALVL
SWT183_1430046	AVAAVDSSNNRASFSSTGSAVEIAAPGVNILSTTPGNSYASVNGTSMASPHVAGVAALVW
B_sp_BAD02409.1	AVAAVDSSNNRASFSSTGPAVEIAAPGVNILSTTPGNSYASVNGTSMASPHVAGVAALVL
B_aurantiacus_WP_026690432.1	AVAAVDQNNRATFSSTGPAVEISAPGVSIILSTTPGNYYAAFNGTSMASPHVAGVAAQVW
B_akibai_WP_035661169.1	AVAAVDQNNRATFSSTGPAVEISAPGVSIILSTIPNNGYASVNGTSMASPHVAGVAALLL
B_pseudofirmus_OF4_ADC50469	AVAAVDQNNRATFSSTGPAVEISAPGAGIILSTTPNNNYASFNGTSMASPHVAGVAAQVW
B_spLL01_WP_047973355.1	AVAAVDQNNRATFSSTGPAVEISAPGVGILSTTPNNNYASFNGTSMASPHVAGVAAQVW
B_marmarensis_WP_022626565_ERN55058.1	AVAAIDSNRRANFSSHGPAVEISAPGVSVLSTTPGNNYASFNGTSMASPHVAGVAALIK
A_transvaalensis_WP_026475840.1	AVGAVDQNNRASFSSHGPAVELAAPGVGVLSVTPGNGYSSLNGTSMASPHVAGVAAQVW
H_halophilus_WP_027963976.1	AVGATDQNNRASFQYGTGLNIVAPGVVQSTYPGNTYASLNGTSMATPHVAGAAALVK
B_lentus_P29600	AVGATDQNNRASFQYGTGLNIVAPGVVQSTYPGNRYASLSGTSMATPHVAGVAALVK
B_sp_Sendai_BAA06157	AVGAVDQNGRANFSSYGVSELEIMAPGVNINSTYLNNGYRSVNGTSMASPHVAGVAALVK
B_pseudofirmus_ADC49870	AVGAVDSSNQSRASFSSVGPPELDVMAPGVSIQSTLPGNKYGAVNGTSMASPHVAGAAALIL
B_amyloliquefaciens_CAA24990	AVGAVNSSNQSRASFSSAGSELDVMAPGVSIQSTLPGGTYGAYNGTSMATPHVAGAAALIL
B_subtilis_WP_015383267	AVGAVDSNSNRASFSSVGAELVMAPGAGVYSTYPTNTYATLNGTSMASPHVAGAAALIL
B_licheniformis_CAJ70731.1	AVGAVDSNNRRASFSSVGSQLEVMAPGVNILSTLPGNSYGSVNGTSMASPHVAGAAALLL
B_sp_KSM-LD1_SB_BAD21128	AVGAVDSNNRASFSSVGSSELEVMAPGVSIILSTVPGSSYASVNGTSMASPHVAGAAALLK
B_sp_KSM-LD1_SA_BAD11988	AVGAVDSSNNRASFSSVGSSELEVMAPGVSIILSTTPGNNYSSFNNGTSMASPHVAGAAALIK
B_sp_m3_13_WP_010192403	AVGAVSSNNTRASFSSVGSSELEVMAPGVNILSTTPGNNYASFNGTSMASPHVAGAAALIK
Bacillus_sp_SprC_AAC43580	AVGAVSSNNTRASFSSVGSSELEVMAPGVNILSTTPGNNYASFNGTSMASPHVAGAAALIK

**FIG. 3D**

BspE04637	AANPNLSNVELRNRLNDTAQNLGDANHFNGNLYRAVDAINGT
SWT183_1430046	AANPNLSNVELRNRLNDTAQNLGDANHFHGLYRAVDAINGTSSGDNGGGDDGGSGPTKP
B_sp_BAD02409.1	AANPNLSNVELRNRLNDTAQNLGDANHFNGNLYRAVDAINGTSSGDNGGGDDGGSEPTKP
B_aurantiacus_WP_026690432.1	QAKPELSNVELRNLLNETAVNLGRSNQFGHGLVQSLDAIQQ
B_akibai_WP_035661169.1	ENNSNLNTELRRELLQSSAKSLGTASQYGYGLVQAMDAINQ
B_pseudofirmus_OF4_ADC50469	QAKPHLSNVELRNLLNDTALPLGGSNQFGNGLVQSMAAIQQ
B_spLL01_WP_047973355.1	QAKPHLSNVALRNLLNETAINLGSSTQYGNGLVQSLDAIQQ
B_marmarensis_WP_022626565_ERN55058.1	QAKPHLSNVELRNLLNDTALPLGGSNQFGNGLVQSMAAIQQ
A_transvaalensis_WP_026475840.1	GANPTLTNSQIRQIMNDTALNLGTWNYNGNLYRAMNGINLALTY
H_halophilus_WP_027963976.1	EAKPHLSNVQLRSLQQAQNLGNSNYGSGLVKSYQAITH
B_lentus_P29600	QKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR
B_sp_Sendai_BAA06157	QKNPSWSNTQIRQHLTSTATSLGNSNQFGSGLVNAEAATR
B_pseudofirmus_ADC49870	QKHPHLTAAQIRNRMNQTAIPLGNSTYYGNGLVDAEYAAQ
B_amyloliquefaciens_CAA24990	SKHPNWTNTQVRSSLENTTKLGDsfyYgKGLINVQAAAQ
B_subtilis_WP_015383267	SKHPTWTNAQVRDRLESTATYLGNSFYYGKGLINVQAAAQ
B_licheniformis_CAJ70731.1	SKHPNLSASQVRNRLSSTATYLGSSFFYYGKGLINVEAAAQ
B_sp_KSM-LD1_SB_BAD21128	AQDPTLTNVQVREILRDTATNLGSSFFYYGNGLVIDVEKALQ
B_sp_KSM-LD1_SA_BAD11988	AKYPNWSAAQIRNKLNSSTTYLGSSFFYYGNGLINVERALQ
B_sp_m3_13_WP_010192403	AKYPSMTNVQIREKLNKNTATNLGDAFYGHGVINVESALQ
Bacillus_sp_SprC_AAC43580	AKYPSMTNVQIRERLNKNTATNLGDPFFYgKGLINVESALQ

**FIG. 3E**

BspE04637	-----	SEQ ID NO:11
SWT183_1430046	GNGKGNKGN	SEQ ID NO:12
B_sp_BAD02409.1	GNGKGNKGN	SEQ ID NO:13
B_aurantiacus_WP_026690432.1	-----	SEQ ID NO:14
B_akibai_WP_035661169.1	-----	SEQ ID NO:15
B_pseudofirmus_OF4_ADC50469	-----	SEQ ID NO:16
B_spLL01_WP_047973355.1	-----	SEQ ID NO:17
B_marmarensis_WP_022626565_ERN55058.1	-----	SEQ ID NO:18
A_transvaalensis_WP_026475840.1	-----	SEQ ID NO:19
H_halophilus_WP_027963976.1	-----	SEQ ID NO:20
B_lentus_P29600	-----	SEQ ID NO:21
B_sp_Sendai_BAA06157	-----	SEQ ID NO:22
B_pseudofirmus_ADC49870	-----	SEQ ID NO:23
B_amyloliquefaciens_CAA24990	-----	SEQ ID NO:24
B_subtilis_WP_015383267	-----	SEQ ID NO:25
B_licheniformis_CAJ70731.1	-----	SEQ ID NO:26
B_sp_KSM-LD1_SB_BAD21128	-----	SEQ ID NO:27
B_sp_KSM-LD1_SA_BAD11988	-----	SEQ ID NO:28
B_sp_m3_13_WP_010192403	-----	SEQ ID NO:29
Bacillus_sp_SprC_AAC43580	-----	SEQ ID NO:30

**FIG. 3F**

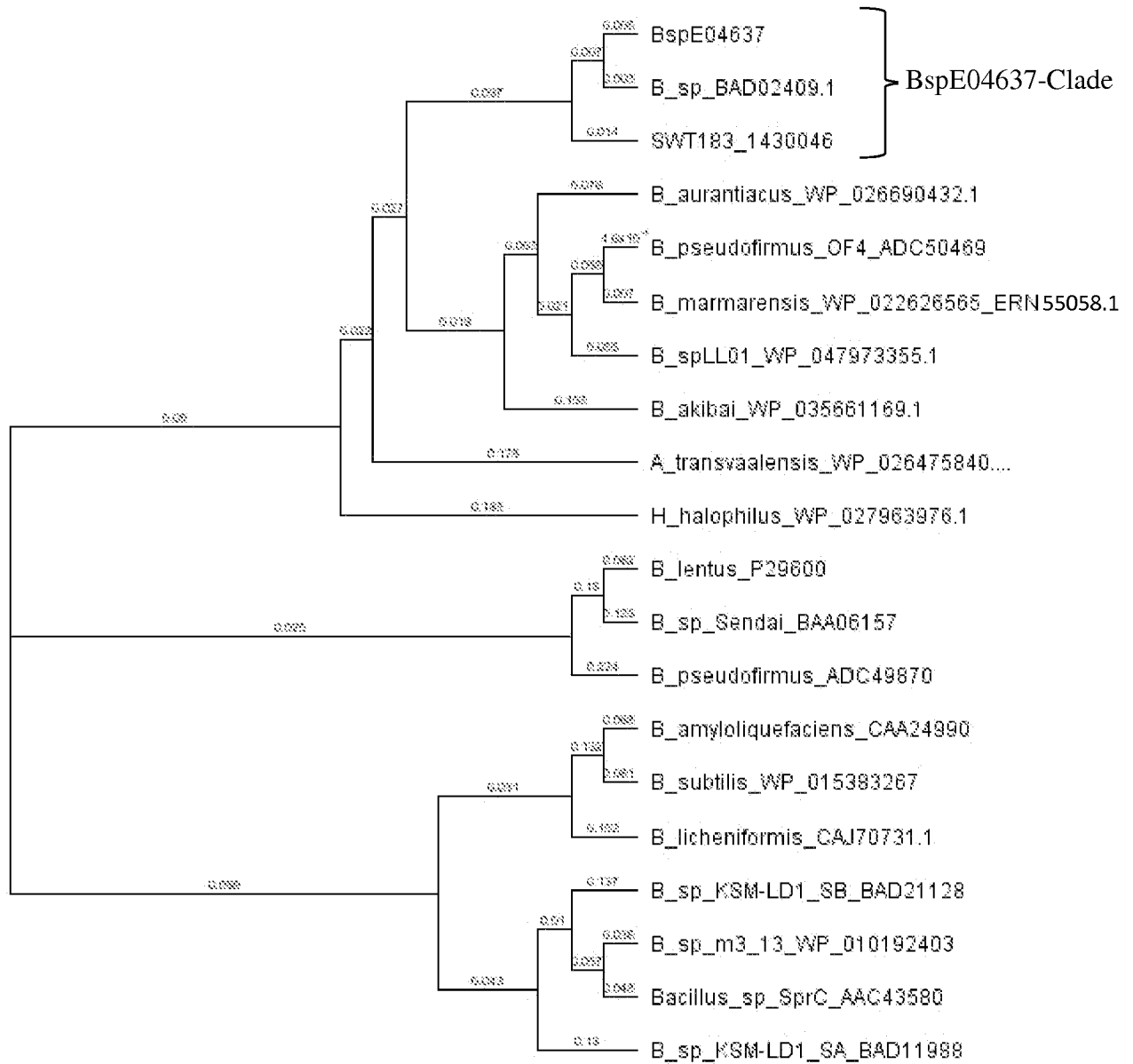
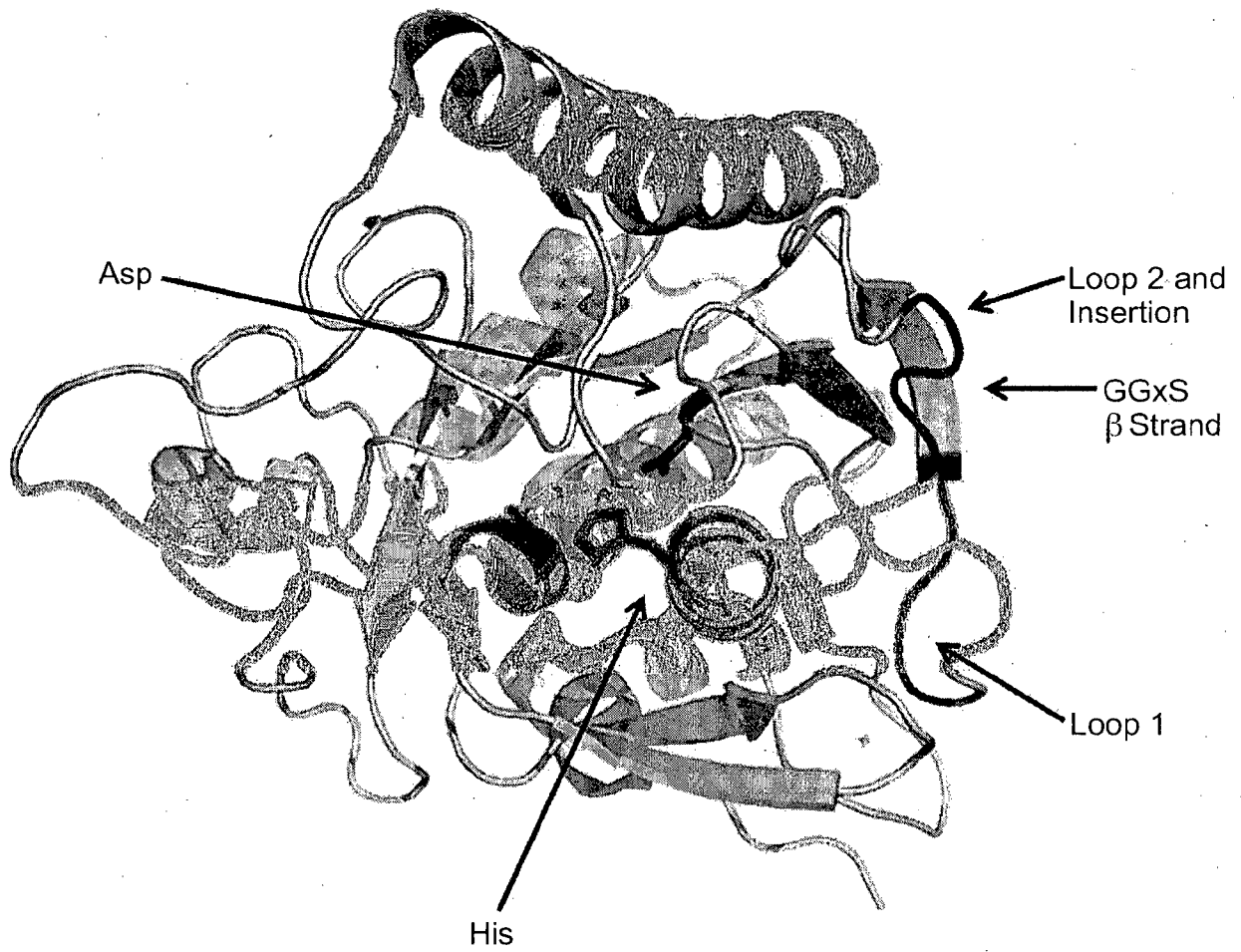


Figure 4





**FIG. 5**

# INTERNATIONAL SEARCH REPORT

International application No PCT/US2015/057512
---

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> INV. C12N9/54                      C11D3/386 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) C12N C11D		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, BIOSIS, Sequence Search, EMBASE, FSTA, WPI Data		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2003 325186 A (KAO CORP) 18 November 2003 (2003-11-18) cited in the application sequence 1 claims 1-6 paragraphs [0021], [0028], [0035] - [0036] figures 1-2 -& DATABASE Geneseq [Online]	1-38
	15 June 2007 (2007-06-15), "Bacillus sp strain KSM-LD1 alkaline protease.", XP002755169, retrieved from EBI accession no. GSP:ADF53282 Database accession no. ADF53282 sequence	
	-/--	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.		
<input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents :		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search  <div style="text-align: center; font-weight: bold;">8 March 2016</div>	Date of mailing of the international search report  <div style="text-align: center; font-weight: bold;">23/03/2016</div>	
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer  <div style="text-align: center; font-weight: bold;">Behrens, Joyce</div>	

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2015/057512

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	-& DATABASE EMBL [Online]  12 June 2008 (2008-06-12), "Alkali protease.", XP002755170, retrieved from EBI accession no. EM_PAT:DJ429568 Database accession no. DJ429568 sequence  -----	
X	DATABASE UniProt [Online]  5 July 2004 (2004-07-05), "SubName: Full=Protease {ECO:0000313 EMBL:BAD02409.1}";", XP002754211, retrieved from EBI accession no. UNIPROT:Q76L29 Database accession no. Q76L29 the whole document  -----	1-10
X	DATABASE UniProt [Online]  19 March 2014 (2014-03-19), "SubName: Full=Uncharacterized protein {ECO:0000313 EMBL:GAE33250.1}";", XP002754212, retrieved from EBI accession no. UNIPROT:W4QM98 Database accession no. W4QM98 the whole document  -----	6,8-10, 19
A	WO 2010/056634 A1 (DANISCO US INC [US]; ESTELL DAVID A [US]; GOEDEGEBUUR FRITS [US]; POUL) 20 May 2010 (2010-05-20) claims 1-9  -----	1-38
X,P	DATABASE UniParc [Online]  18 October 2015 (2015-10-18), "UPI0006D1704D", XP002754213, retrieved from EBI accession no. UNIPARC:UPI0006D1704D Database accession no. UPI0006D1704D sequence -& DATABASE NCBI RefSeq [Online]  15 October 2015 (2015-10-15), "SubName: peptidase S8 [Piscibacillus salipiscarius]", XP002755171, retrieved from NCBI accession no. WP_054754483.1 Database accession no. WP_054754483.1 sequence  -----	1-10

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No PCT/US2015/057512
---

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
JP 2003325186 A	18-11-2003	JP 4179797 B2 JP 2003325186 A	12-11-2008 18-11-2003
-----			
WO 2010056634 A1	20-05-2010	AR 074104 A1 AR 074105 A1 BR PI0920925 A2 CA 2742992 A1 CA 2743060 A1 CN 102209782 A CN 105154418 A EP 2362902 A1 EP 2362903 A1 JP 5412522 B2 JP 5508431 B2 JP 2012508029 A JP 2012508317 A RU 2011123745 A US 2010152088 A1 US 2011281328 A1 US 2014295527 A1 US 2015240192 A1 WO 2010056634 A1 WO 2010056635 A1	22-12-2010 22-12-2010 18-08-2015 20-05-2010 20-05-2010 05-10-2011 16-12-2015 07-09-2011 07-09-2011 12-02-2014 28-05-2014 05-04-2012 05-04-2012 20-12-2012 17-06-2010 17-11-2011 02-10-2014 27-08-2015 20-05-2010 20-05-2010
-----			