

(19) **DANMARK**



Patent- og
Varemærkestyrelsen

(10) **DK/EP 2702153 T3**

(12) Oversættelse af
europæisk patentskrift

-
- (51) Int.Cl.: **C 12 N 9/42 (2006.01)**
- (45) Oversættelsen bekendtgjort den: **2019-03-18**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2018-12-12**
- (86) Europæisk ansøgning nr.: **12723981.2**
- (86) Europæisk indleveringsdag: **2012-04-26**
- (87) Den europæiske ansøgnings publiceringsdag: **2014-03-05**
- (86) International ansøgning nr.: **US2012035257**
- (87) Internationalt publikationsnr.: **WO2012149192**
- (30) Prioritet: **2011-04-28 US 201161480245 P**
- (84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**
- (73) Patenthaver: **Novozymes, Inc., 1445 Drew Avenue, Davis, CA 95618, USA**
Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Danmark
- (72) Opfinder: **MCBRAYER, Brett, 3712 Kos Island Avenue, Sacramento, California 95834, USA**
SPODSBERG, Nikolaj, Vandkarsevej 62, 2880 Bagsværd, Danmark
CASSLAND, Pierre, Lövsångaregatan 9, 235 37 Vellinge, Sverige
- (54) Benævnelse: **Polypeptider med endoglucanaseaktivitet og polynukleotider, som koder for dem**
- (56) Fremdragne publikationer:
WO-A1-2010/096931
WO-A1-2010/141325
WO-A2-2008/088724
WO-A2-2010/118257
GB-A- 2 159 834
JP-A- 2011 182 677
US-A- 5 047 332
US-A- 5 861 271
P FÄHNRIK ET AL: "AFFINITY CHROMATOGRAPHY OF EXTRACELLULAR CELLULASE FROM CHAETOMIUM-CELLULOLYTICUM", BIOTECHNOLOGY LETTERS, 1 January 1984 (1984-01-01), pages 251-256, XP55034947, DOI: 10.1007/BF00140046
DATABASE UniProt [Online] 16 November 2011 (2011-11-16), "SubName: Full=Glycoside hydrolase family 5 protein; EC=3.2.1.4;", XP002681493, retrieved from EBI accession no. UNIPROT:G2RD75 Database accession no. G2RD75
H DUBEAU ET AL: "XYLANASE OF CHAETOMIUM-CELLULOLYTICUM ITS NATURE OF PRODUCTION AND HYDROLYTIC POTENTIAL", BIOTECHNOLOGY LETTERS, vol. 9, no. 4, 1 January 1987 (1987-01-01) , pages 275-280, XP55034948, DOI: 10.1007/bf01027163
ANKUDIMOVA N V ET AL: "Cellulase complex from Chaetomium cellulolyticum: isolation and properties of

Fortsættes ...

major components.", BIOCHEMISTRY. BIOKHIMIJA SEP 1999, vol. 64, no. 9, September 1999 (1999-09), pages 1068-1073, ISSN: 0006-2979

DESCRIPTION

Background of the Invention

Field of the Invention

[0001] The present invention relates to polypeptides having endoglucanase activity, catalytic domains, and cellulose binding domains, and polynucleotides encoding the polypeptides, catalytic domains, and cellulose binding domains. The invention also relates to nucleic acid constructs, vectors, and host cells comprising the polynucleotides as well as methods of producing and using the polypeptides, catalytic domains, and cellulose binding domains.

Description of the Related Art

[0002] Cellulose is a polymer of the simple sugar glucose covalently linked by beta-1,4-bonds. Many microorganisms produce enzymes that hydrolyze beta-linked glucans. These enzymes include endoglucanases, cellobiohydrolases, and beta-glucosidases. Endoglucanases digest the cellulose polymer at random locations, opening it to attack by cellobiohydrolases. Cellobiohydrolases sequentially release molecules of cellobiose from the ends of the cellulose polymer. Cellobiose is a water-soluble beta-1,4-linked dimer of glucose. Beta-glucosidases hydrolyze cellobiose to glucose.

[0003] The conversion of lignocellulosic feedstocks into ethanol has the advantages of the ready availability of large amounts of feedstock, the desirability of avoiding burning or land filling the materials, and the cleanliness of the ethanol fuel. Wood, agricultural residues, herbaceous crops, and municipal solid wastes have been considered as feedstocks for ethanol production. These materials primarily consist of cellulose, hemicellulose, and lignin. Once the lignocellulose is converted to fermentable sugars, e.g., glucose, the fermentable sugars are easily fermented by yeast into ethanol.

[0004] The present invention provides polypeptides having endoglucanase activity and polynucleotides encoding the polypeptides. US5,861,271 discloses an endoglucanase II catalytic domain which shares 81% identity with the endoglucanase of the present invention. Fährnich and Irrgang, 1984, Biotechnology Letters vol. 6:4, p. 251-256 discloses fractionation of fermentation broth from *C. cellulolyticum* ATCC32319 to separate fractions having CMC activity from beta-glucosidase activity.

Summary of the Invention

[0005] The present invention relates to isolated polypeptides having endoglucanase activity selected from the group consisting of:

1. (a) a polypeptide having at least 90% sequence identity to amino acids 18 to 419 of SEQ ID NO: 2.

[0006] The present invention also relates to isolated polynucleotides encoding the polypeptides of the present invention; nucleic acid constructs; recombinant expression vectors; recombinant host cells comprising the polynucleotides; and methods of producing the polypeptides.

[0007] The present invention also relates to a composition comprising the polypeptide of the invention.

[0008] The present invention also relates to methods for degrading or converting a cellulosic material, comprising: treating the cellulosic material with an enzyme composition of the present invention.

[0009] The present invention also relates to methods of producing a fermentation product, comprising: (a) saccharifying a cellulosic material with an enzyme composition of the present invention; (b) fermenting the saccharified cellulosic material with one or more (e.g., several) fermenting microorganisms to produce the fermentation product; and (c) recovering the fermentation product from the fermentation.

Brief Description of the Figures

[0010]

Figure 1 shows a restriction map of pDAU222-PE03860005710.

Figure 2 shows the effect of *Chaetomium virescens* GH5 endoglucanase on a high-temperature enzyme composition in the hydrolysis of milled unwashed PCS at 50-65°C.

Figure 3 shows the effect of *C. virescens* GH5 endoglucanase on handsheet tensile index of unbleached eucalyptus Kraft pulp.

Figure 4 shows the effect of *C. virescens* GH5 endoglucanase on handsheet tear index of unbleached eucalyptus Kraft pulp.

Figure 5 shows a comparison of the effect of *Chaetomium virescens* GH5 endoglucanase and *Trichoderma reesei* GH5 endoglucanase II in the hydrolysis of milled unwashed PCS by a cellulase enzyme composition.

Definitions

[0011] Acetylxyylan esterase: The term "acetylxyylan esterase" means a carboxylesterase (EC 3.1.1.72) that catalyzes the hydrolysis of acetyl groups from polymeric xylan, acetylated xylose, acetylated glucose, alpha-naphthyl acetate, and p-nitrophenyl acetate. For purposes of the present invention, acetylxyylan esterase activity is determined using 0.5 mM p-nitrophenylacetate as substrate in 50 mM sodium acetate pH 5.0 containing 0.01% TWEEN™ 20 (polyoxyethylene sorbitan monolaurate). One unit of acetylxyylan esterase is defined as the amount of enzyme capable of releasing 1 μ mole of p-nitrophenolate anion per minute at pH 5, 25°C.

[0012] Allelic variant: The term "allelic variant" means any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequences. An allelic variant of a polypeptide is a polypeptide encoded by an allelic variant of a gene.

[0013] Alpha-L-arabinofuranosidase: The term "alpha-L-arabinofuranosidase" means an alpha-L-arabinofuranoside arabinofuranohydrolase (EC 3.2.1.55) that catalyzes the hydrolysis of terminal non-reducing alpha-L-arabinofuranoside residues in alpha-L-arabinosides. The enzyme acts on alpha-L-arabinofuranosides, alpha-L-arabinans containing (1,3)- and/or (1,5)-linkages, arabinoxylans, and arabinogalactans. Alpha-L-arabinofuranosidase is also known as arabinosidase, alpha-arabinosidase, alpha-L-arabinosidase, alpha-arabinofuranosidase, polysaccharide alpha-L-arabinofuranosidase, alpha-L-arabinofuranoside hydrolase, L-arabinosidase, or alpha-L-arabinanase. For purposes of the present invention, alpha-L-arabinofuranosidase activity is determined using 5 mg of medium viscosity wheat arabinoxylan (Megazyme International Ireland, Ltd., Bray, Co. Wicklow, Ireland) per ml of 100 mM sodium acetate pH 5 in a total volume of 200 μ l for 30 minutes at 40°C followed by arabinose analysis by AMINEX® HPX-87H column chromatography (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

[0014] Alpha-glucuronidase: The term "alpha-glucuronidase" means an alpha-D-glucosiduronate glucuronohydrolase (EC 3.2.1.139) that catalyzes the hydrolysis of an alpha-D-glucuronoside to D-glucuronate and an alcohol. For purposes of the present invention, alpha-glucuronidase activity is determined according to de Vries, 1998, J. Bacteriol. 180: 243-249. One unit of alpha-glucuronidase equals the amount of enzyme capable of releasing 1 μ mole of glucuronic or 4-O-methylglucuronic acid per minute at pH 5, 40°C.

[0015] Beta-glucosidase: The term "beta-glucosidase" means a beta-D-glucoside glucohydrolase (E.C. 3.2.1.21) that catalyzes the hydrolysis of terminal non-reducing beta-D-

glucose residues with the release of beta-D-glucose. For purposes of the present invention, beta-glucosidase activity is determined using p-nitrophenyl-beta-D-glucopyranoside as substrate according to the procedure of Venturi et al., 2002, Extracellular beta-D-glucosidase from *Chaetomium thermophilum* var. *coprophilum*: production, purification and some biochemical properties, J. Basic Microbiol. 42: 55-66. One unit of beta-glucosidase is defined as 1.0 μ mole of p-nitrophenolate anion produced per minute at 25°C, pH 4.8 from 1 mM p-nitrophenyl-beta-D-glucopyranoside as substrate in 50 mM sodium citrate containing 0.01% TWEEN® 20.

[0016] Beta-xylosidase: The term "beta-xylosidase" means a beta-D-xyloside xylohydrolase (E.C. 3.2.1.37) that catalyzes the exo-hydrolysis of short beta (1→4)-xylooligosaccharides to remove successive D-xylose residues from non-reducing termini. For purposes of the present invention, one unit of beta-xylosidase is defined as 1.0 μ mole of p-nitrophenolate anion produced per minute at 40°C, pH 5 from 1 mM p-nitrophenyl-beta-D-xyloside as substrate in 100 mM sodium citrate containing 0.01% TWEEN® 20.

[0017] Catalytic domain: The term "catalytic domain" means the region of an enzyme containing the catalytic machinery of the enzyme.

[0018] cDNA: The term "cDNA" means a DNA molecule that can be prepared by reverse transcription from a mature, spliced, mRNA molecule obtained from a eukaryotic or prokaryotic cell. cDNA lacks intron sequences that may be present in the corresponding genomic DNA. The initial, primary RNA transcript is a precursor to mRNA that is processed through a series of steps, including splicing, before appearing as mature spliced mRNA.

[0019] Cellobiohydrolase: The term "cellobiohydrolase" means a 1,4-beta-D-glucan cellobiohydrolase (E.C. 3.2.1.91 and E.C. 3.2.1.176) that catalyzes the hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, cellooligosaccharides, or any beta-1,4-linked glucose containing polymer, releasing cellobiose from the reducing or non-reducing ends of the chain (Teeri, 1997, Crystalline cellulose degradation: New insight into the function of cellobiohydrolases, Trends in Biotechnology 15: 160-167; Teeri et al., 1998, Trichoderma reesei cellobiohydrolases: why so efficient on crystalline cellulose?, Biochem. Soc. Trans. 26: 173-178). Cellobiohydrolase activity is determined according to the procedures described by Lever et al., 1972, Anal. Biochem. 47: 273-279; van Tilbeurgh et al., 1982, FEBS Letters, 149: 152-156; van Tilbeurgh and Claeysens, 1985, FEBS Letters, 187: 283-288; and Tomme et al., 1988, Eur. J. Biochem. 170: 575-581. In the present invention, the Tomme *et al.* method can be used to determine cellobiohydrolase activity.

[0020] Cellulose binding domain: The term "cellulose binding domain" means the region of an enzyme that mediates binding of the enzyme to amorphous regions of a cellulose substrate. The cellulose binding domain (CBD) is typically found either at the N-terminal or at the C-terminal extremity of an endoglucanase. The term "cellulose binding domain" is also known in the art as a "carbohydrate binding module".

[0021] Cellulolytic enzyme or cellulase: The term "cellulolytic enzyme" or "cellulase" means one or more (*e.g.*, several) enzymes that hydrolyze a cellulosic material. Such enzymes include endoglucanase(s), cellobiohydrolase(s), beta-glucosidase(s), or combinations thereof. The two basic approaches for measuring cellulolytic activity include: (1) measuring the total cellulolytic activity, and (2) measuring the individual cellulolytic activities (endoglucanases, cellobiohydrolases, and beta-glucosidases) as reviewed in Zhang et al., Outlook for cellulase improvement: Screening and selection strategies, 2006, *Biotechnology Advances* 24: 452-481. Total cellulolytic activity is usually measured using insoluble substrates, including Whatman N°1 filter paper, microcrystalline cellulose, bacterial cellulose, algal cellulose, cotton, pretreated lignocellulose, *etc.* The most common total cellulolytic activity assay is the filter paper assay using Whatman N°1 filter paper as the substrate. The assay was established by the International Union of Pure and Applied Chemistry (IUPAC) (Ghose, 1987, Measurement of cellulase activities, *Pure Appl. Chem.* 59: 257-68).

[0022] For purposes of the present invention, cellulolytic enzyme activity is determined by measuring the increase in hydrolysis of a cellulosic material by cellulolytic enzyme(s) under the following conditions: 1-50 mg of cellulolytic enzyme protein/g of cellulose in PCS (or other pretreated cellulosic material) for 3-7 days at a suitable temperature, *e.g.*, 50°C, 55°C, or 60°C, compared to a control hydrolysis without addition of cellulolytic enzyme protein. Typical conditions are 1 ml reactions, washed or unwashed PCS, 5% insoluble solids, 50 mM sodium acetate pH 5, 1 mM MnSO₄, 50°C, 55°C, or 60°C, 72 hours, sugar analysis by AMINEX® HPX-87H column (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

[0023] Cellulosic material: The term "cellulosic material" means any material containing cellulose. The predominant polysaccharide in the primary cell wall of biomass is cellulose, the second most abundant is hemicellulose, and the third is pectin. The secondary cell wall, produced after the cell has stopped growing, also contains polysaccharides and is strengthened by polymeric lignin covalently cross-linked to hemicellulose. Cellulose is a homopolymer of anhydrocellobiose and thus a linear beta-(1-4)-D-glucan, while hemicelluloses include a variety of compounds, such as xylans, xyloglucans, arabinoxylans, and mannans in complex branched structures with a spectrum of substituents. Although generally polymorphous, cellulose is found in plant tissue primarily as an insoluble crystalline matrix of parallel glucan chains. Hemicelluloses usually hydrogen bond to cellulose, as well as to other hemicelluloses, which help stabilize the cell wall matrix.

[0024] Cellulose is generally found, for example, in the stems, leaves, hulls, husks, and cobs of plants or leaves, branches, and wood of trees. The cellulosic material can be, but is not limited to, agricultural residue, herbaceous material (including energy crops), municipal solid waste, pulp and paper mill residue, waste paper, and wood (including forestry residue) (see, for example, Wiselogle et al., 1995, in *Handbook on Bioethanol* (Charles E. Wyman, editor), pp.105-118, Taylor & Francis, Washington D.C.; Wyman, 1994, *Bioresource Technology* 50: 3-16; Lynd, 1990, *Applied Biochemistry and Biotechnology* 24/25: 695-719; Mosier et al., 1999, Recent Progress in Bioconversion of Lignocellulosics, in *Advances in Biochemical Engineering/Biotechnology*, T. Scheper, managing editor, Volume 65, pp.23-40, Springer-

Verlag, New York). It is understood herein that the cellulose may be in the form of lignocellulose, a plant cell wall material containing lignin, cellulose, and hemicellulose in a mixed matrix. In a preferred aspect, the cellulosic material is any biomass material. In another preferred aspect, the cellulosic material is lignocellulose, which comprises cellulose, hemicelluloses, and lignin.

[0025] In one aspect, the cellulosic material is agricultural residue. In another aspect, the cellulosic material is herbaceous material (including energy crops). In another aspect, the cellulosic material is municipal solid waste. In another aspect, the cellulosic material is pulp and paper mill residue. In another aspect, the cellulosic material is waste paper. In another aspect, the cellulosic material is wood (including forestry residue).

[0026] In another aspect, the cellulosic material is arundo. In another aspect, the cellulosic material is bagasse. In another aspect, the cellulosic material is bamboo. In another aspect, the cellulosic material is corn cob. In another aspect, the cellulosic material is corn fiber. In another aspect, the cellulosic material is corn stover. In another aspect, the cellulosic material is miscanthus. In another aspect, the cellulosic material is orange peel. In another aspect, the cellulosic material is rice straw. In another aspect, the cellulosic material is switchgrass. In another aspect, the cellulosic material is wheat straw.

[0027] In another aspect, the cellulosic material is aspen. In another aspect, the cellulosic material is eucalyptus. In another aspect, the cellulosic material is fir. In another aspect, the cellulosic material is pine. In another aspect, the cellulosic material is poplar. In another aspect, the cellulosic material is spruce. In another aspect, the cellulosic material is willow.

[0028] In another aspect, the cellulosic material is algal cellulose. In another aspect, the cellulosic material is bacterial cellulose. In another aspect, the cellulosic material is cotton linter. In another aspect, the cellulosic material is filter paper. In another aspect, the cellulosic material is microcrystalline cellulose. In another aspect, the cellulosic material is phosphoric-acid treated cellulose.

[0029] In another aspect, the cellulosic material is an aquatic biomass. As used herein the term "aquatic biomass" means biomass produced in an aquatic environment by a photosynthesis process. The aquatic biomass can be algae, emergent plants, floating-leaf plants, or submerged plants.

[0030] The cellulosic material may be used as is or may be subjected to pretreatment, using conventional methods known in the art, as described herein. In a preferred aspect, the cellulosic material is pretreated.

[0031] Coding sequence: The term "coding sequence" means a polynucleotide, which directly specifies the amino acid sequence of a polypeptide. The boundaries of the coding sequence are generally determined by an open reading frame, which begins with a start codon such as ATG, GTG, or TTG and ends with a stop codon such as TAA, TAG, or TGA. The

coding sequence may be a genomic DNA, cDNA, synthetic DNA, or a combination thereof.

[0032] Control sequences: The term "control sequences" means nucleic acid sequences necessary for expression of a polynucleotide encoding a mature polypeptide of the present invention. Each control sequence may be native (*i.e.*, from the same gene) or foreign (*i.e.*, from a different gene) to the polynucleotide encoding the polypeptide or native or foreign to each other. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the polynucleotide encoding a polypeptide.

[0033] Endoglucanase: The term "endoglucanase" means an endo-1,4-(1,3;1,4)-beta-D-glucan 4-glucanohydrolase (E.C. 3.2.1.4) that catalyzes endohydrolysis of 1,4-beta-D-glycosidic linkages in cellulose, cellulose derivatives (such as carboxymethyl cellulose and hydroxyethyl cellulose), lichenin, beta-1,4 bonds in mixed beta-1,3 glucans such as cereal beta-D-glucans or xyloglucans, and other plant material containing cellulosic components. Endoglucanase activity can be determined by measuring reduction in substrate viscosity or increase in reducing ends determined by a reducing sugar assay (Zhang et al., 2006, *Biotechnology Advances* 24: 452-481). For purposes of the present invention, endoglucanase activity is determined using carboxymethyl cellulose (CMC) as substrate according to the procedure of Ghose, 1987, *Pure and Appl. Chem.* 59: 257-268, at pH 5, 40°C.

[0034] In one aspect, the polypeptides of the present invention have at least 20%, e.g., at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the endoglucanase activity of the mature polypeptide of SEQ ID NO: 2.

[0035] Expression: The term "expression" includes any step involved in the production of a polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

[0036] Expression vector: The term "expression vector" means a linear or circular DNA molecule that comprises a polynucleotide encoding a polypeptide and is operably linked to control sequences that provide for its expression.

[0037] Family 61 glycoside hydrolase: The term "Family 61 glycoside hydrolase" or "Family GH61" or "GH61" means a polypeptide falling into the glycoside hydrolase Family 61 according to Henrissat B., 1991, A classification of glycosyl hydrolases based on amino-acid sequence similarities, *Biochem. J.* 280: 309-316, and Henrissat B., and Bairoch A., 1996, Updating the sequence-based classification of glycosyl hydrolases, *Biochem. J.* 316: 695-696. The enzymes in this family were originally classified as a glycoside hydrolase family based on measurement of very weak endo-1,4-beta-D-glucanase activity in one family member. The structure and mode of action of these enzymes are non-canonical and they cannot be considered as bona

fide glycosidases. However, they are kept in the CAZy classification on the basis of their capacity to enhance the breakdown of lignocellulose when used in conjunction with a cellulase or a mixture of cellulases.

[0038] Feruloyl esterase: The term "feruloyl esterase" means a 4-hydroxy-3-methoxycinnamoyl-sugar hydrolase (EC 3.1.1.73) that catalyzes the hydrolysis of 4-hydroxy-3-methoxycinnamoyl (feruloyl) groups from esterified sugar, which is usually arabinose in natural biomass substrates, to produce ferulate (4-hydroxy-3-methoxycinnamate). Feruloyl esterase is also known as ferulic acid esterase, hydroxycinnamoyl esterase, FAE-III, cinnamoyl ester hydrolase, FAEA, cinnAE, FAE-I, or FAE-II. For purposes of the present invention, feruloyl esterase activity is determined using 0.5 mM *p*-nitrophenylferulate as substrate in 50 mM sodium acetate pH 5.0. One unit of feruloyl esterase equals the amount of enzyme capable of releasing 1 μ mole of *p*-nitrophenolate anion per minute at pH 5, 25°C.

[0039] Fragment: The term "fragment" means a polypeptide or a catalytic or cellulose binding domain having one or more (e.g., several) amino acids absent from the amino and/or carboxyl terminus of a mature polypeptide or domain; wherein the fragment has endoglucanase. In one aspect, a fragment having endoglucanase activity contains at least 340 amino acid residues, e.g., at least 360 amino acid residues or at least 380 amino acid residues. In another aspect, a fragment of a catalytic domain having endoglucanase activity contains at least 280 amino acid residues, e.g., at least 295 amino acid residues or at least 310 amino acid residues. In another aspect, a fragment of a cellulose binding domain contains at least 30 amino acids, e.g., at least 32 amino acid residues or at least 34 amino acid residues.

[0040] Hemicellulolytic enzyme or hemicellulase: The term "hemicellulolytic enzyme" or "hemicellulase" means one or more (e.g., several) enzymes that hydrolyze a hemicellulosic material. See, for example, Shallom, D. and Shoham, Y. Microbial hemicellulases. Current Opinion In Microbiology, 2003, 6(3): 219-228). Hemicellulases are key components in the degradation of plant biomass. Examples of hemicellulases include, but are not limited to, an acetylmannan esterase, an acetylxylan esterase, an arabinanase, an arabinofuranosidase, a coumaric acid esterase, a feruloyl esterase, a galactosidase, a glucuronidase, a glucuronoyl esterase, a mannanase, a mannosidase, a xylanase, and a xylosidase. The substrates of these enzymes, the hemicelluloses, are a heterogeneous group of branched and linear polysaccharides that are bound via hydrogen bonds to the cellulose microfibrils in the plant cell wall, crosslinking them into a robust network. Hemicelluloses are also covalently attached to lignin, forming together with cellulose a highly complex structure. The variable structure and organization of hemicelluloses require the concerted action of many enzymes for its complete degradation. The catalytic modules of hemicellulases are either glycoside hydrolases (GHs) that hydrolyze glycosidic bonds, or carbohydrate esterases (CEs), which hydrolyze ester linkages of acetate or ferulic acid side groups. These catalytic modules, based on homology of their primary sequence, can be assigned into GH and CE families. Some families, with an overall similar fold, can be further grouped into clans, marked alphabetically (e.g., GH-A). A most informative and updated classification of these and other carbohydrate active enzymes is available in the Carbohydrate-Active Enzymes (CAZy) database. Hemicellulolytic enzyme

activities can be measured according to Ghose and Bisaria, 1987, Pure & Appl. Chem. 59: 1739-1752, at a suitable temperature, e.g., 50°C, 55°C, or 60°C, and pH, e.g., 5.0 or 5.5.

[0041] High stringency conditions: The term "high stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 50% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 65°C.

[0042] Host cell: The term "host cell" means any cell type that is susceptible to transformation, transfection, transduction, or the like with a nucleic acid construct or expression vector comprising a polynucleotide of the present invention. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

[0043] Isolated: The term "isolated" means a substance in a form or environment that does not occur in nature. Non-limiting examples of isolated substances include (1) any non-naturally occurring substance, (2) any substance including, but not limited to, any enzyme, variant, nucleic acid, protein, peptide or cofactor, that is at least partially removed from one or more or all of the naturally occurring constituents with which it is associated in nature; (3) any substance modified by the hand of man relative to that substance found in nature; or (4) any substance modified by increasing the amount of the substance relative to other components with which it is naturally associated (e.g., recombinant production in a host cell; multiple copies of a gene encoding the substance; and use of a stronger promoter than the promoter naturally associated with the gene encoding the substance).

[0044] Low stringency conditions: The term "low stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 25% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 50°C.

[0045] Mature polypeptide: The term "mature polypeptide" means a polypeptide in its final form following translation and any post-translational modifications, such as N-terminal processing, C-terminal truncation, glycosylation, phosphorylation, etc. In one aspect, the mature polypeptide is amino acids 18 to 419 of SEQ ID NO: 2 based on the SignalP program (Nielsen et al., 1997, Protein Engineering 10: 1-6) that predicts amino acids 1 to 17 of SEQ ID NO: 2 are a signal peptide. It is known in the art that a host cell may produce a mixture of two or more different mature polypeptides (*i.e.*, with a different C-terminal and/or N-terminal amino acid) expressed by the same polynucleotide.

[0046] Mature polypeptide coding sequence: The term "mature polypeptide coding sequence" means a polynucleotide that encodes a mature polypeptide having endoglucanase activity. In one aspect, the mature polypeptide coding sequence is nucleotides 52 to 1317 of

SEQ ID NO: 1 or the cDNA sequence thereof based on the SignalP program (Nielsen *et al.*, 1997, *supra*) that predicts nucleotides 1 to 51 of SEQ ID NO: 1 encode a signal peptide.

[0047] Medium stringency conditions: The term "medium stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 35% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 55°C.

[0048] Medium-high stringency conditions: The term "medium-high stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 35% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 60°C.

[0049] Nucleic acid construct: The term "nucleic acid construct" means a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature or which is synthetic, which comprises one or more control sequences.

[0050] Operably linked: The term "operably linked" means a configuration in which a control sequence is placed at an appropriate position relative to the coding sequence of a polynucleotide such that the control sequence directs expression of the coding sequence.

[0051] Polypeptide having cellulolytic enhancing activity: The term "polypeptide having cellulolytic enhancing activity" means a GH61 polypeptide that catalyzes the enhancement of the hydrolysis of a cellulosic material by enzyme having cellulolytic activity. For purposes of the present invention, cellulolytic enhancing activity is determined by measuring the increase in reducing sugars or the increase of the total of cellobiose and glucose from the hydrolysis of a cellulosic material by cellulolytic enzyme under the following conditions: 1-50 mg of total protein/g of cellulose in PCS, wherein total protein is comprised of 50-99.5% w/w cellulolytic enzyme protein and 0.5-50% w/w protein of a GH61 polypeptide having cellulolytic enhancing activity for 1-7 days at a suitable temperature, *e.g.*, 50°C, 55°C, or 60°C, and pH, *e.g.*, 5.0 or 5.5, compared to a control hydrolysis with equal total protein loading without cellulolytic enhancing activity (1-50 mg of cellulolytic protein/g of cellulose in PCS). In a preferred aspect, a mixture of CELLUCLAST® 1.5L (Novozymes A/S, Bagsvaerd, Denmark) in the presence of 2-3% of total protein weight *Aspergillus oryzae* beta-glucosidase (recombinantly produced in *Aspergillus oryzae* according to WO 02/095014) or 2-3% of total protein weight *Aspergillus fumigatus* beta-glucosidase (recombinantly produced in *Aspergillus oryzae* as described in WO 2002/095014) of cellulase protein loading is used as the source of the cellulolytic activity.

[0052] The GH61 polypeptides having cellulolytic enhancing activity enhance the hydrolysis of a cellulosic material catalyzed by enzyme having cellulolytic activity by reducing the amount of

cellulolytic enzyme required to reach the same degree of hydrolysis preferably at least 1.01-fold, e.g., at least 1.05-fold, at least 1.10-fold, at least 1.25-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 10-fold, or at least 20-fold.

[0053] Pretreated corn stover: The term "PCS" or "Pretreated Corn Stover" means a cellulosic material derived from corn stover by treatment with heat and dilute sulfuric acid, alkaline pretreatment, or neutral pretreatment.

[0054] Sequence identity: The relatedness between two amino acid sequences or between two nucleotide sequences is described by the parameter "sequence identity".

[0055] For purposes of the present invention, the sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends Genet. 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

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

[0056] For purposes of the present invention, the sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *supra*) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, *supra*), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

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

[0057] Subsequence: The term "subsequence" means a polynucleotide having one or more (e.g., several) nucleotides absent from the 5' and/or 3' end of a mature polypeptide coding sequence; wherein the subsequence encodes a fragment having endoglucanase activity. In one aspect, a subsequence contains at least 1020 nucleotides, e.g., at least 1080 nucleotides or at least 1140 nucleotides. In another aspect, a subsequence of a coding sequence of a catalytic domain having endoglucanase activity contains at least 840 nucleotides, e.g., at least 885 nucleotides or at least 930 nucleotides. In another aspect, a subsequence of a coding sequence of a cellulose binding domain contains at least 90 nucleotides, e.g., at least 96 nucleotides or at least 102 nucleotides.

[0058] Variant: The term "variant" means a polypeptide having endoglucanase activity comprising an alteration, *i.e.*, a substitution, insertion, and/or deletion, at one or more (*e.g.*, several) positions. A substitution means replacement of the amino acid occupying a position with a different amino acid; a deletion means removal of the amino acid occupying a position; and an insertion means adding an amino acid adjacent to and immediately following the amino acid occupying a position.

[0059] Very high stringency conditions: The term "very high stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 50% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 70°C.

[0060] Very low stringency conditions: The term "very low stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 25% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 45°C.

[0061] Xylan-containing material: The term "xylan-containing material" means any material comprising a plant cell wall polysaccharide containing a backbone of beta-(1-4)-linked xylose residues. Xylans of terrestrial plants are heteropolymers possessing a beta-(1-4)-D-xylopyranose backbone, which is branched by short carbohydrate chains. They comprise D-glucuronic acid or its 4-O-methyl ether, L-arabinose, and/or various oligosaccharides, composed of D-xylose, L-arabinose, D- or L-galactose, and D-glucose. Xylan-type polysaccharides can be divided into homoxylans and heteroxylans, which include glucuronoxylans, (arabino)glucuronoxylans, (glucurono)arabinoxylans, arabinoxylans, and complex heteroxylans. See, for example, Ebringerova et al., 2005, *Adv. Polym. Sci.* 186: 1-67.

[0062] In the methods of the present invention, any material containing xylan may be used. In a preferred aspect, the xylan-containing material is lignocellulose.

[0063] Xylan degrading activity or xylanolytic activity: The term "xylan degrading activity" or "xylanolytic activity" means a biological activity that hydrolyzes xylan-containing material. The two basic approaches for measuring xylanolytic activity include: (1) measuring the total xylanolytic activity, and (2) measuring the individual xylanolytic activities (*e.g.*, endoxylanases, beta-xylosidases, arabinofuranosidases, alpha-glucuronidases, acetylxylan esterases, feruloyl esterases, and alpha-glucuronoyl esterases). Recent progress in assays of xylanolytic enzymes was summarized in several publications including Biely and Puchard, Recent progress in the assays of xylanolytic enzymes, 2006, *Journal of the Science of Food and Agriculture* 86(11): 1636-1647; Spanikova and Biely, 2006, Glucuronoyl esterase - Novel carbohydrate esterase produced by *Schizophyllum commune*, *FEBS Letters* 580(19): 4597-4601; Herrmann, Vrsanska, Jurickova, Hirsch, Biely, and Kubicek, 1997, The beta-D-xylosidase of *Trichoderma*

reesei is a multifunctional beta-D-xylan xylohydrolase, Biochemical Journal 321: 375-381.

[0064] Total xylan degrading activity can be measured by determining the reducing sugars formed from various types of xylan, including, for example, oat spelt, beechwood, and larchwood xylans, or by photometric determination of dyed xylan fragments released from various covalently dyed xylans. The most common total xylanolytic activity assay is based on production of reducing sugars from polymeric 4-O-methyl glucuronoxylan as described in Bailey, Biely, Poutanen, 1992, Interlaboratory testing of methods for assay of xylanase activity, Journal of Biotechnology 23(3): 257-270. Xylanase activity can also be determined with 0.2% AZCL-arabinoxylan as substrate in 0.01% TRITON® X-100 (4-(1,1,3,3-tetramethylbutyl)phenyl-polyethylene glycol) and 200 mM sodium phosphate buffer pH 6 at 37°C. One unit of xylanase activity is defined as 1.0 µmole of azurine produced per minute at 37°C, pH 6 from 0.2% AZCL-arabinoxylan as substrate in 200 mM sodium phosphate pH 6 buffer.

[0065] For purposes of the present invention, xylan degrading activity is determined by measuring the increase in hydrolysis of birchwood xylan (Sigma Chemical Co., Inc., St. Louis, MO, USA) by xylan-degrading enzyme(s) under the following typical conditions: 1 ml reactions, 5 mg/ml substrate (total solids), 5 mg of xylanolytic protein/g of substrate, 50 mM sodium acetate pH 5, 50°C, 24 hours, sugar analysis using p-hydroxybenzoic acid hydrazide (PHBAH) assay as described by Lever, 1972, A new reaction for colorimetric determination of carbohydrates, Anal. Biochem 47: 273-279.

[0066] Xylanase: The term "xylanase" means a 1,4-beta-D-xylan-xylohydrolase (E.C. 3.2.1.8) that catalyzes the endohydrolysis of 1,4-beta-D-xylosidic linkages in xylans. For purposes of the present invention, xylanase activity is determined with 0.2% AZCL-arabinoxylan as substrate in 0.01% TRITON® X-100 and 200 mM sodium phosphate buffer pH 6 at 37°C. One unit of xylanase activity is defined as 1.0 µmole of azurine produced per minute at 37°C, pH 6 from 0.2% AZCL-arabinoxylan as substrate in 200 mM sodium phosphate pH 6 buffer.

Detailed Description of the Invention

Polypeptides Having Endoglucanase Activity

[0067] In an aspect, the present disclosure relates to isolated polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, which have endoglucanase activity. In one aspect, the polypeptides differ by up to 10 amino acids, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10, from the mature polypeptide of SEQ ID NO: 2.

[0068] A polypeptide of the present invention preferably comprises or consists of amino acids 18 to 419 of SEQ ID NO: 2.

[0069] In another embodiment, the present disclosure relates to isolated polypeptides having endoglucanase activity encoded by polynucleotides that hybridize under very low stringency conditions, low stringency conditions, medium stringency conditions, medium-high stringency conditions, high stringency conditions, or very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, (ii) the cDNA sequence thereof, or (iii) the full-length complement of (i) or (ii) (Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, New York).

[0070] The polynucleotide of SEQ ID NO: 1 or a subsequence thereof, as well as the polypeptide of SEQ ID NO: 2 or a fragment thereof, may be used to design nucleic acid probes to identify and clone DNA encoding polypeptides having endoglucanase activity from strains of different genera or species according to methods well known in the art. In particular, such probes can be used for hybridization with the genomic DNA or cDNA of a cell of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Such probes can be considerably shorter than the entire sequence, but should be at least 15, e.g., at least 25, at least 35, or at least 70 nucleotides in length. Preferably, the nucleic acid probe is at least 100 nucleotides in length, e.g., at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, at least 500 nucleotides, at least 600 nucleotides, at least 700 nucleotides, at least 800 nucleotides, or at least 900 nucleotides in length. Both DNA and RNA probes can be used. The probes are typically labeled for detecting the corresponding gene (for example, with ^{32}P , ^3H , ^{35}S , biotin, or avidin). Such probes are encompassed by the present invention.

[0071] A genomic DNA or cDNA library prepared from such other strains may be screened for DNA that hybridizes with the probes described above and encodes a polypeptide having endoglucanase activity. Genomic or other DNA from such other strains may be separated by agarose or polyacrylamide gel electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to and immobilized on nitrocellulose or other suitable carrier material. In order to identify a clone or DNA that hybridizes with SEQ ID NO: 1 or a subsequence thereof, the carrier material is used in a Southern blot.

[0072] For purposes of the present disclosure, hybridization indicates that the polynucleotide hybridizes to a labeled nucleic acid probe corresponding to (i) SEQ ID NO: 1; (ii) the mature polypeptide coding sequence of SEQ ID NO: 1; (iii) the cDNA sequence thereof; (iv) the full-length complement thereof; or (v) a subsequence thereof; under very low to very high stringency conditions. Molecules to which the nucleic acid probe hybridizes under these conditions can be detected using, for example, X-ray film or any other detection means known in the art.

[0073] In one aspect, the nucleic acid probe is nucleotides 52 to 1317 of SEQ ID NO: 1. In another aspect, the nucleic acid probe is a polynucleotide that encodes the polypeptide of SEQ ID NO: 2; the mature polypeptide thereof; or a fragment thereof. In another aspect, the nucleic acid probe is SEQ ID NO: 1 or the cDNA sequence thereof.

[0074] In another aspect, the present disclosure relates to isolated polypeptides having endoglucanase activity encoded by polynucleotides having a sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1, or the cDNA sequence thereof, of at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%.

[0075] In another aspect, the present disclosure relates to variants of the mature polypeptide of SEQ ID NO: 2 comprising a substitution, deletion, and/or insertion at one or more (e.g., several) positions. In an aspect, the number of amino acid substitutions, deletions and/or insertions introduced into the mature polypeptide of SEQ ID NO: 2 is up to 10, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10. The amino acid changes may be of a minor nature, that is conservative amino acid substitutions or insertions that do not significantly affect the folding and/or activity of the protein; small deletions, typically of 1-30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to 20-25 residues; or a small extension that facilitates purification by changing net charge or another function, such as a poly-histidine tract, an antigenic epitope or a binding domain.

[0076] Examples of conservative substitutions are within the groups of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine and asparagine), hydrophobic amino acids (leucine, isoleucine and valine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine, serine, threonine and methionine). Amino acid substitutions that do not generally alter specific activity are known in the art and are described, for example, by H. Neurath and R.L. Hill, 1979, In, *The Proteins*, Academic Press, New York. Common substitutions are Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, and Asp/Gly.

[0077] Alternatively, the amino acid changes are of such a nature that the physico-chemical properties of the polypeptides are altered. For example, amino acid changes may improve the thermal stability of the polypeptide, alter the substrate specificity, change the pH optimum, and the like.

[0078] Essential amino acids in a polypeptide can be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, 1989, *Science* 244: 1081-1085). In the latter technique, single alanine mutations are introduced at every residue in the molecule, and the resultant mutant molecules are tested for endoglucanase activity to identify amino acid residues that are critical to the activity of the molecule. See also, Hilton et al., 1996, *J. Biol. Chem.* 271: 4699-4708. The active site of the enzyme or other biological interaction can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic resonance, crystallography, electron diffraction, or photoaffinity labeling, in conjunction with mutation of putative contact site amino acids. See, for example, de Vos et al., 1992, *Science* 255: 306-312; Smith et al., 1992, *J. Mol. Biol.* 224: 899-904; Wlodaver et al., 1992, *FEBS Lett.* 309: 59-64. The identity of essential

amino acids can also be inferred from an alignment with a related polypeptide.

[0079] Single or multiple amino acid substitutions, deletions, and/or insertions can be made and tested using known methods of mutagenesis, recombination, and/or shuffling, followed by a relevant screening procedure, such as those disclosed by Reidhaar-Olson and Sauer, 1988, *Science* 241: 53-57; Bowie and Sauer, 1989, *Proc. Natl. Acad. Sci. USA* 86: 2152-2156; WO 95/17413; or WO 95/22625. Other methods that can be used include error-prone PCR, phage display (e.g., Lowman et al., 1991, *Biochemistry* 30: 10832-10837; U.S. Patent No. 5,223,409; WO 92/06204), and region-directed mutagenesis (Derbyshire et al., 1986, *Gene* 46: 145; Ner et al., 1988, *DNA* 7: 127).

[0080] Mutagenesis/shuffling methods can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides expressed by host cells (Ness et al., 1999, *Nature Biotechnology* 17: 893-896). Mutagenized DNA molecules that encode active polypeptides can be recovered from the host cells and rapidly sequenced using standard methods in the art. These methods allow the rapid determination of the importance of individual amino acid residues in a polypeptide.

[0081] The polypeptide may be a hybrid polypeptide in which a region of one polypeptide is fused at the N-terminus or the C-terminus of a region of another polypeptide.

[0082] The polypeptide may be a fusion polypeptide or cleavable fusion polypeptide in which another polypeptide is fused at the N-terminus or the C-terminus of the polypeptide of the present invention. A fusion polypeptide is produced by fusing a polynucleotide encoding another polypeptide to a polynucleotide of the present invention. Techniques for producing fusion polypeptides are known in the art, and include ligating the coding sequences encoding the polypeptides so that they are in frame and that expression of the fusion polypeptide is under control of the same promoter(s) and terminator. Fusion polypeptides may also be constructed using intein technology in which fusion polypeptides are created post-translationally (Cooper et al., 1993, *EMBO J.* 12: 2575-2583; Dawson et al., 1994, *Science* 266: 776-779).

[0083] A fusion polypeptide can further comprise a cleavage site between the two polypeptides. Upon secretion of the fusion protein, the site is cleaved releasing the two polypeptides. Examples of cleavage sites include, but are not limited to, the sites disclosed in Martin et al., 2003, *J. Ind. Microbiol. Biotechnol.* 3: 568-576; Svetina et al., 2000, *J. Biotechnol.* 76: 245-251; Rasmussen-Wilson et al., 1997, *Appl. Environ. Microbiol.* 63: 3488-3493; Ward et al., 1995, *Biotechnology* 13: 498-503; and Contreras et al., 1991, *Biotechnology* 9: 378-381; Eaton et al., 1986, *Biochemistry* 25: 505-512; Collins-Racie et al., 1995, *Biotechnology* 13: 982-987; Carter et al., 1989, *Proteins: Structure, Function, and Genetics* 6: 240-248; and Stevens, 2003, *Drug Discovery World* 4: 35-48.

Sources of Polypeptides Having Endoglucanase Activity

[0084] A polypeptide having endoglucanase activity of the present invention may be obtained from microorganisms of any genus.

[0085] The polypeptide may be a fungal polypeptide.

[0086] In another aspect, the polypeptide is a *Chaetomium virescens* polypeptide, e.g., a polypeptide obtained from *Chaetomium virescens* ATCC 32319.

[0087] Strains of these species are readily accessible to the public in a number of culture collections, such as the American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau Voor Schimmelcultures (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

[0088] The polypeptide may be identified and obtained from other sources including microorganisms isolated from nature (e.g., soil, composts, water, etc.) or DNA samples obtained directly from natural materials (e.g., soil, composts, water, etc.) using the above-mentioned probes. Techniques for isolating microorganisms and DNA directly from natural habitats are well known in the art. A polynucleotide encoding the polypeptide may then be obtained by similarly screening a genomic DNA or cDNA library of another microorganism or mixed DNA sample. Once a polynucleotide encoding a polypeptide has been detected with the probe(s), the polynucleotide can be isolated or cloned by utilizing techniques that are known to those of ordinary skill in the art (see, e.g., Sambrook *et al.*, 1989, *supra*).

Catalytic Domains

[0089] In one aspect, the present disclosure also relates to catalytic domains having a sequence identity to amino acids 18 to 419 of SEQ ID NO: 2 of at least 85%, e.g., at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%. In one aspect, the catalytic domains comprise amino acid sequences that differ by up to 10 amino acids, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10, from amino acids 93 to 419 of SEQ ID NO: 2.

[0090] The catalytic domain preferably comprises or consists of amino acids 93 to 419 of SEQ ID NO: 2 or an allelic variant thereof; or is a fragment thereof having endoglucanase activity.

[0091] In another aspect, the present disclosure also relates to catalytic domains encoded by polynucleotides that hybridize under very low stringency conditions, low stringency conditions, medium stringency conditions, medium-high stringency conditions, high stringency conditions, or very high stringency conditions (as defined above) with (i) the nucleotides 277 to 1317 of SEQ ID NO: 1, (ii) the cDNA sequence thereof, or (iii) the full-length complement of (i) or (ii) (Sambrook *et al.*, 1989, *supra*).

[0092] In another embodiment, the present disclosure also relates to catalytic domains encoded by polynucleotides having a sequence identity to nucleotides 277 to 1317 of SEQ ID NO: 1 or the cDNA sequence thereof, of at least 85%, *e.g.*, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%.

[0093] The polynucleotide encoding the catalytic domain preferably comprises or consists of nucleotides 277 to 1317 of SEQ ID NO: 1.

[0094] In another aspect, the present disclosure also relates to catalytic domain variants of amino acids 93 to 419 of SEQ ID NO: 2 comprising a substitution, deletion, and/or insertion at one or more (*e.g.*, several) positions. In one aspect, the number of amino acid substitutions, deletions and/or insertions introduced into the sequence of amino acids 93 to 419 of SEQ ID NO: 2 is up to 10, *e.g.*, 1, 2, 3, 4, 5, 6, 8, 9, or 10.

Cellulose Binding Domains

[0095] In one aspect, the present disclosure also relates to cellulose binding domains having a sequence identity to amino acids 23 to 58 of SEQ ID NO: 2 of at least 80%, *e.g.*, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%. In one aspect, the cellulose binding domains comprise amino acid sequences that differ by up to 10 amino acids, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10, from amino acids 23 to 58 of SEQ ID NO: 2.

[0096] The cellulose binding domain preferably comprises or consists of amino acids 23 to 58 of SEQ ID NO: 2 or an allelic variant thereof; or is a fragment thereof having cellulose binding activity.

[0097] In another aspect, the present disclosure also relates to cellulose binding domains encoded by polynucleotides that hybridize under very low stringency conditions, low stringency conditions, medium stringency conditions, medium-high stringency conditions, high stringency conditions, or very high stringency conditions (as defined above) with nucleotides 67 to 174 of SEQ ID NO: 1 or the full-length complement thereof (Sambrook *et al.*, 1989, *supra*).

[0098] In another aspect, the present disclosure also relates to cellulose binding domains encoded by polynucleotides having a sequence identity to nucleotides 67 to 174 of SEQ ID NO: 1 of at least 80%, *e.g.*, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%.

[0099] The polynucleotide encoding the cellulose binding domain preferably comprises or consists of nucleotides 67 to 174 of SEQ ID NO: 1.

[0100] In another aspect, the present disclosure also relates to cellulose binding domain variants of amino acids 23 to 58 of SEQ ID NO: 2 comprising a substitution, deletion, and/or insertion at one or more (e.g., several) positions. In one aspect, the number of amino acid substitutions, deletions and/or insertions introduced into the sequence of amino acids 23 to 58 of SEQ ID NO: 2 is up to 10, e.g., 1, 2, 3, 4, 5, 6, 8, 9, or 10.

[0101] A catalytic domain operably linked to the cellulose binding domain may be from a hydrolase, isomerase, ligase, lyase, oxidoreductase, or transferase, e.g., an aminopeptidase, amylase, carbohydrase, carboxypeptidase, catalase, cellobiohydrolase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, endoglucanase, esterase, alpha-galactosidase, beta-galactosidase, glucoamylase, alpha-glucosidase, beta-glucosidase, invertase, laccase, lipase, mannosidase, mutanase, oxidase, pectinolytic enzyme, peroxidase, phytase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transglutaminase, xylanase, or beta-xylosidase. The polynucleotide encoding the catalytic domain may be obtained from any prokaryotic, eukaryotic, or other source.

Polynucleotides

[0102] The present invention also relates to isolated polynucleotides encoding a polypeptide, of the present invention, as described herein.

[0103] The techniques used to isolate or clone a polynucleotide encoding a polypeptide are known in the art and include isolation from genomic DNA or cDNA, or a combination thereof. The cloning of the polynucleotides from genomic DNA can be effected, e.g., by using the well known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, e.g., Innis et al., 1990, PCR: A Guide to Methods and Application, Academic Press, New York. Other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligation activated transcription (LAT) and polynucleotide-based amplification (NASBA) may be used. The polynucleotides may be cloned from a strain of *Chaetomium*, or a related organism and thus, for example, may be an allelic or species variant of the polypeptide encoding region of the polynucleotide.

[0104] Modification of a polynucleotide encoding a polypeptide of the present invention may be necessary for synthesizing polypeptides substantially similar to the polypeptide. The term "substantially similar" to the polypeptide refers to non-naturally occurring forms of the polypeptide. These polypeptides may differ in some engineered way from the polypeptide isolated from its native source, e.g., variants that differ in specific activity, thermostability, pH optimum, or the like. The variants may be constructed on the basis of the polynucleotide presented as the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof, e.g., a subsequence thereof, and/or by introduction of nucleotide substitutions that do

not result in a change in the amino acid sequence of the polypeptide, but which correspond to the codon usage of the host organism intended for production of the enzyme, or by introduction of nucleotide substitutions that may give rise to a different amino acid sequence. For a general description of nucleotide substitution, see, e.g., Ford et al., 1991, Protein Expression and Purification 2: 95-107.

Nucleic Acid Constructs

[0105] The present invention also relates to nucleic acid constructs comprising a polynucleotide of the present invention operably linked to one or more control sequences that direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences.

[0106] The polynucleotide may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the polynucleotide prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides utilizing recombinant DNA methods are well known in the art.

[0107] The control sequence may be a promoter, a polynucleotide that is recognized by a host cell for expression of a polynucleotide encoding a polypeptide of the present invention. The promoter contains transcriptional control sequences that mediate the expression of the polypeptide. The promoter may be any polynucleotide that shows transcriptional activity in the host cell including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

[0108] Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a bacterial host cell are the promoters obtained from the *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus subtilis* *xyIA* and *xyIB* genes, *Bacillus thuringiensis* *cryIIIA* gene (Agaisse and Lereclus, 1994, Molecular Microbiology 13: 97-107), *E. coli lac* operon, *E. coli trc* promoter (Egon et al., 1988, Gene 69: 301-315), *Streptomyces coelicolor* agarase gene (*dagA*), and prokaryotic beta-lactamase gene (Villa-Kamaroff et al., 1978, Proc. Natl. Acad. Sci. USA 75: 3727-3731), as well as the *tac* promoter (DeBoer et al., 1983, Proc. Natl. Acad. Sci. USA 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in Gilbert et al., 1980, Scientific American 242: 74-94; and in Sambrook et al., 1989, *supra*. Examples of tandem promoters are disclosed in WO 99/43835.

[0109] Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a filamentous fungal host cell are promoters obtained from the

genes for *Aspergillus nidulans* acetamidase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Aspergillus oryzae* TAKA amylase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Fusarium oxysporum* trypsin-like protease (WO 96/00787), *Fusarium venenatum* amyloglucosidase (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), *Rhizomucor miehei* lipase, *Rhizomucor miehei* aspartic proteinase, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei* xylanase II, *Trichoderma reesei* xylanase III, *Trichoderma reesei* beta-xylosidase, and *Trichoderma reesei* translation elongation factor, as well as the NA2-tpi promoter (a modified promoter from an *Aspergillus* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus* triose phosphate isomerase gene; non-limiting examples include modified promoters from an *Aspergillus niger* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus nidulans* or *Aspergillus oryzae* triose phosphate isomerase gene); and mutant, truncated, and hybrid promoters thereof. Other promoters are described in U.S. Patent No. 6,011,147.

[0110] In a yeast host, useful promoters are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH1, ADH2/GAP), *Saccharomyces cerevisiae* triose phosphate isomerase (TPI), *Saccharomyces cerevisiae* metallothionein (CUP1), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase. Other useful promoters for yeast host cells are described by Romanos et al., 1992, Yeast 8: 423-488.

[0111] The control sequence may also be a transcription terminator, which is recognized by a host cell to terminate transcription. The terminator is operably linked to the 3'-terminus of the polynucleotide encoding the polypeptide. Any terminator that is functional in the host cell may be used in the present invention.

[0112] Preferred terminators for bacterial host cells are obtained from the genes for *Bacillus clausii* alkaline protease (*aprH*), *Bacillus licheniformis* alpha-amylase (*amyL*), and *Escherichia coli* ribosomal RNA (*rrnB*).

[0113] Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* acetamidase, *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase, *Aspergillus oryzae* TAKA amylase, *Fusarium oxysporum* trypsin-like protease, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei*

xylanase II, *Trichoderma reesei* xylanase III, *Trichoderma reesei* beta-xylosidase, and *Trichoderma reesei* translation elongation factor.

[0114] Preferred terminators for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase, *Saccharomyces cerevisiae* cytochrome C (CYC1), and *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos *et al.*, 1992, *supra*.

[0115] The control sequence may also be an mRNA stabilizer region downstream of a promoter and upstream of the coding sequence of a gene which increases expression of the gene.

[0116] Examples of suitable mRNA stabilizer regions are obtained from a *Bacillus thuringiensis cryIIIA* gene (WO 94/25612) and a *Bacillus subtilis* SP82 gene (Hue *et al.*, 1995, *Journal of Bacteriology* 177: 3465-3471).

[0117] The control sequence may also be a leader, a nontranslated region of an mRNA that is important for translation by the host cell. The leader is operably linked to the 5'-terminus of the polynucleotide encoding the polypeptide. Any leader that is functional in the host cell may be used.

[0118] Preferred leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase.

[0119] Suitable leaders for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, *Saccharomyces cerevisiae* alpha-factor, and *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP).

[0120] The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3'-terminus of the polynucleotide and, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence that is functional in the host cell may be used.

[0121] Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase *Aspergillus oryzae* TAKA amylase, and *Fusarium oxysporum* trypsin-like protease.

[0122] Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, *Mol. Cellular Biol.* 15: 5983-5990.

[0123] The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a polypeptide and directs the polypeptide into the cell's

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

[0124] Effective signal peptide coding sequences for bacterial host cells are the signal peptide coding sequences obtained from the genes for *Bacillus* NCIB 11837 maltogenic amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus stearothermophilus* neutral proteases (*nprT*, *nprS*, *nprM*), and *Bacillus subtilis* *prsA*. Further signal peptides are described by Simonen and Palva, 1993, Microbiological Reviews 57: 109-137.

[0125] Effective signal peptide coding sequences for filamentous fungal host cells are the signal peptide coding sequences obtained from the genes for *Aspergillus niger* neutral amylase, *Aspergillus niger* glucoamylase, *Aspergillus oryzae* TAKA amylase, *Humicola insolens* cellulase, *Humicola insolens* endoglucanase V, *Humicola lanuginosa* lipase, and *Rhizomucor miehei* aspartic proteinase.

[0126] Useful signal peptides for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase. Other useful signal peptide coding sequences are described by Romanos *et al.*, 1992, *supra*.

[0127] The control sequence may also be a propeptide coding sequence that encodes a propeptide positioned at the N-terminus of a polypeptide. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to an active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding sequence may be obtained from the genes for *Bacillus subtilis* alkaline protease (*aprE*), *Bacillus subtilis* neutral protease (*nprT*), *Myceliophthora thermophila* laccase (WO 95/33836), *Rhizomucor miehei* aspartic proteinase, and *Saccharomyces cerevisiae* alpha-factor.

[0128] Where both signal peptide and propeptide sequences are present, the propeptide sequence is positioned next to the N-terminus of a polypeptide and the signal peptide sequence is positioned next to the N-terminus of the propeptide sequence.

It may also be desirable to add regulatory sequences that regulate expression of the polypeptide relative to the growth of the host cell. Examples of regulatory sequences are those that cause expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory sequences in

prokaryotic systems include the *lac*, *tac*, and *trp* operator systems. In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the *Aspergillus niger* glucoamylase promoter, *Aspergillus oryzae* TKA alpha-amylase promoter, and *Aspergillus oryzae* glucoamylase promoter, *Trichoderma reesei* cellobiohydrolase I promoter, and *Trichoderma reesei* cellobiohydrolase II promoter may be used. Other examples of regulatory sequences are those that allow for gene amplification. In eukaryotic systems, these regulatory sequences include the dihydrofolate reductase gene that is amplified in the presence of methotrexate, and the metallothionein genes that are amplified with heavy metals. In these cases, the polynucleotide encoding the polypeptide would be operably linked to the regulatory sequence.

Expression Vectors

[0129] The present invention also relates to recombinant expression vectors comprising a polynucleotide of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleotide and control sequences may be joined together to produce a recombinant expression vector that may include one or more convenient restriction sites to allow for insertion or substitution of the polynucleotide encoding the polypeptide at such sites. Alternatively, the polynucleotide may be expressed by inserting the polynucleotide or a nucleic acid construct comprising the polynucleotide into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

[0130] The recombinant expression vector may be any vector (e.g., a plasmid or virus) that can be conveniently subjected to recombinant DNA procedures and can bring about expression of the polynucleotide. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector may be a linear or closed circular plasmid.

[0131] The vector may be an autonomously replicating vector, i.e., a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one that, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids that together contain the total DNA to be introduced into the genome of the host cell, or a transposon, may be used.

[0132] The vector preferably contains one or more selectable markers that permit easy selection of transformed, transfected, transduced, or the like cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

[0133] Examples of bacterial selectable markers are *Bacillus licheniformis* or *Bacillus subtilis*

dal genes, or markers that confer antibiotic resistance such as ampicillin, chloramphenicol, kanamycin, neomycin, spectinomycin, or tetracycline resistance. Suitable markers for yeast host cells include, but are not limited to, ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. Selectable markers for use in a filamentous fungal host cell include, but are not limited to, *adeA* (phosphoribosylaminoimidazole-succinocarboxamide synthase), *adeB* (phosphoribosylaminoimidazole synthase), *amdS* (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hph* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenylyltransferase), and *trpC* (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are *Aspergillus nidulans* or *Aspergillus oryzae* *amdS* and *pyrG* genes and a *Streptomyces hygroscopicus* *bar* gene. Preferred for use in a *Trichoderma* cell are *adeA*, *adeB*, *amdS*, *hph*, and *pyrG* genes.

[0134] The selectable marker may be a dual selectable marker system as described in WO 2010/039889. In one aspect, the dual selectable marker is a *hph-tk* dual selectable marker system.

[0135] The vector preferably contains an element(s) that permits integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

[0136] For integration into the host cell genome, the vector may rely on the polynucleotide's sequence encoding the polypeptide or any other element of the vector for integration into the genome by homologous or non-homologous recombination. Alternatively, the vector may contain additional polynucleotides for directing integration by homologous recombination into the genome of the host cell at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should contain a sufficient number of nucleic acids, such as 100 to 10,000 base pairs, 400 to 10,000 base pairs, and 800 to 10,000 base pairs, which have a high degree of sequence identity to the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding polynucleotides. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

[0137] For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. The origin of replication may be any plasmid replicator mediating autonomous replication that functions in a cell. The term "origin of replication" or "plasmid replicator" means a polynucleotide that enables a plasmid or vector to replicate *in vivo*.

[0138] Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAM β 1 permitting replication in *Bacillus*.

[0139] Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6.

[0140] Examples of origins of replication useful in a filamentous fungal cell are AMA1 and ANS1 (Gems et al., 1991, *Gene* 98: 61-67; Cullen et al., 1987, *Nucleic Acids Res.* 15: 9163-9175; WO 00/24883). Isolation of the AMA1 gene and construction of plasmids or vectors comprising the gene can be accomplished according to the methods disclosed in WO 00/24883.

[0141] More than one copy of a polynucleotide of the present invention may be inserted into a host cell to increase production of a polypeptide. An increase in the copy number of the polynucleotide can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the polynucleotide where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the polynucleotide, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

[0142] The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, e.g., Sambrook *et al.*, 1989, *supra*).

Host Cells

[0143] The present invention also relates to recombinant host cells, comprising a polynucleotide of the present invention operably linked to one or more control sequences that direct the production of a polypeptide of the present invention. A construct or vector comprising a polynucleotide is introduced into a host cell so that the construct or vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication. The choice of a host cell will to a large extent depend upon the gene encoding the polypeptide and its source.

[0144] The host cell may be any cell useful in the recombinant production of a polypeptide of the present invention, e.g., a prokaryote or a eukaryote.

[0145] The prokaryotic host cell may be any Gram-positive or Gram-negative bacterium. Gram-positive bacteria include, but are not limited to, *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, and *Streptomyces*. Gram-negative bacteria include, but are not limited to, *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, and *Ureaplasma*.

[0146] The bacterial host cell may be any *Bacillus* cell including, but not limited to, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis* cells.

[0147] The bacterial host cell may also be any *Streptococcus* cell including, but not limited to, *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, and *Streptococcus equi* subsp. *Zooepidemicus* cells.

[0148] The bacterial host cell may also be any *Streptomyces* cell including, but not limited to, *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, and *Streptomyces lividans* cells.

[0149] The introduction of DNA into a *Bacillus* cell may be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, Mol. Gen. Genet. 168: 111-115), competent cell transformation (see, e.g., Young and Spizizen, 1961, J. Bacteriol. 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, J. Mol. Biol. 56: 209-221), electroporation (see, e.g., Shigekawa and Dower, 1988, Biotechniques 6: 742-751), or conjugation (see, e.g., Koehler and Thorne, 1987, J. Bacteriol. 169: 5271-5278). The introduction of DNA into an *E. coli* cell may be effected by protoplast transformation (see, e.g., Hanahan, 1983, J. Mol. Biol. 166: 557-580) or electroporation (see, e.g., Dower et al., 1988, Nucleic Acids Res. 16: 6127-6145). The introduction of DNA into a *Streptomyces* cell may be effected by protoplast transformation, electroporation (see, e.g., Gong et al., 2004, Folia Microbiol. (Praha) 49: 399-405), conjugation (see, e.g., Mazodier et al., 1989, J. Bacteriol. 171: 3583-3585), or transduction (see, e.g., Burke et al., 2001, Proc. Natl. Acad. Sci. USA 98: 6289-6294). The introduction of DNA into a *Pseudomonas* cell may be effected by electroporation (see, e.g., Choi et al., 2006, J. Microbiol. Methods 64: 391-397) or conjugation (see, e.g., Pinedo and Smets, 2005, Appl. Environ. Microbiol. 71: 51-57). The introduction of DNA into a *Streptococcus* cell may be effected by natural competence (see, e.g., Perry and Kuramitsu, 1981, Infect. Immun. 32: 1295-1297), protoplast transformation (see, e.g., Catt and Jollick, 1991, Microbios 68: 189-207), electroporation (see, e.g., Buckley et al., 1999, Appl. Environ. Microbiol. 65: 3800-3804), or conjugation (see, e.g., Clewell, 1981, Microbiol. Rev. 45: 409-436). However, any method known in the art for introducing DNA into a host cell can be used.

[0150] The host cell may also be a eukaryote, such as a mammalian, insect, plant, or fungal cell.

[0151] The host cell may be a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota as well as the Oomycota and all mitosporic fungi (as defined by Hawksworth et al., In, Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK).

[0152] The fungal host cell may be a yeast cell. "Yeast" as used herein includes ascosporogenous yeast (Endomycetales), basidiosporogenous yeast, and yeast belonging to the Fungi Imperfecti (Blastomycetes). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in Biology and Activities of Yeast (Skinner, Passmore, and Davenport, editors, Soc. App. Bacteriol. Symposium Series No. 9, 1980).

[0153] The yeast host cell may be a *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* cell such as a *Kluyveromyces lactis*, *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis*, or *Yarrowia lipolytica* cell.

[0154] The fungal host cell may be a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth *et al.*, 1995, *supra*). The filamentous fungi are generally characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

[0155] The filamentous fungal host cell may be an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, or *Trichoderma* cell.

[0156] For example, the filamentous fungal host cell may be an *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Chrysosporium inops*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium merdarium*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium tropicum*, *Chrysosporium zonatum*, *Coprinus cinereus*, *Coriolus hirsutus*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* cell.

[0157] Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* and *Trichoderma* host cells are described in EP 238023, Yelton et al., 1984, Proc. Natl. Acad. Sci. USA 81: 1470-1474, and Christensen et al., 1988, Bio/Technology 6: 1419-1422. Suitable methods for transforming *Fusarium* species are described by Malardier et al., 1989, Gene 78: 147-156, and WO 96/00787. Yeast may be transformed using the procedures described by Becker and Guarente, In Abelson, J.N. and Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito et al., 1983, J. Bacteriol. 153: 163; and Hinnen et al., 1978, Proc. Natl. Acad. Sci. USA 75: 1920.

Methods of Production

[0158] The present disclosure also relates to methods of producing a polypeptide of the present invention, comprising: (a) cultivating a cell, which in its wild-type form produces the polypeptide, under conditions conducive for production of the polypeptide; and optionally (b) recovering the polypeptide. In one aspect, the cell is a *Chaetomium* cell. In another aspect, the cell is a *Chaetomium virescens* cell. In another aspect, the cell is *Chaetomium virescens* ATCC 32319.

[0159] The present invention also relates to methods of producing a polypeptide of the present invention, comprising: (a) cultivating a recombinant host cell of the present invention under conditions conducive for production of the polypeptide; and optionally (b) recovering the polypeptide.

[0160] The host cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods known in the art. For example, the cells may be cultivated by shake flask cultivation, or small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection). If the polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

[0161] The polypeptide may be detected using methods known in the art that are specific for the polypeptides. These detection methods include, but are not limited to, use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme assay may be used to determine the activity of the polypeptide.

[0162] The polypeptide may be recovered using methods known in the art. For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, collection, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation. In one aspect, a fermentation broth comprising the polypeptide is recovered.

[0163] The polypeptide may be purified by a variety of procedures known in the art including, but not limited to, chromatography (e.g., ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing), differential solubility (e.g., ammonium sulfate precipitation), SDS-PAGE, or extraction (see, e.g., Protein Purification, Janson and Ryden, editors, VCH Publishers, New York, 1989) to obtain substantially pure polypeptides.

[0164] In an alternative aspect, the polypeptide is not recovered, but rather a host cell of the present invention expressing the polypeptide is used as a source of the polypeptide.

Plants

[0165] The present invention also relates to isolated plants, e.g., a transgenic plant, plant part, or plant cell, comprising a polynucleotide of the present invention so as to express and produce a polypeptide or domain in recoverable quantities. The polypeptide or domain may be recovered from the plant or plant part. Alternatively, the plant or plant part containing the polypeptide or domain may be used as such for improving the quality of a food or feed, e.g., improving nutritional value, palatability, and rheological properties, or to destroy an antinutritive factor.

[0166] The transgenic plant can be dicotyledonous (a dicot) or monocotyledonous (a monocot). Examples of monocot plants are grasses, such as meadow grass (blue grass, *Poa*), forage grass such as *Festuca*, *Lolium*, temperate grass, such as *Agrostis*, and cereals, e.g., wheat, oats, rye, barley, rice, sorghum, and maize (corn).

[0167] Examples of dicot plants are tobacco, legumes, such as lupins, potato, sugar beet, pea, bean and soybean, and cruciferous plants (family Brassicaceae), such as cauliflower, rape seed, and the closely related model organism *Arabidopsis thaliana*.

[0168] Examples of plant parts are stem, callus, leaves, root, fruits, seeds, and tubers as well as the individual tissues comprising these parts, e.g., epidermis, mesophyll, parenchyme, vascular tissues, meristems. Specific plant cell compartments, such as chloroplasts, apoplasts, mitochondria, vacuoles, peroxisomes and cytoplasm are also considered to be a plant part. Furthermore, any plant cell, whatever the tissue origin, is considered to be a plant part. Likewise, plant parts such as specific tissues and cells isolated to facilitate the utilization of the invention are also considered plant parts, e.g., embryos, endosperms, aleurone and seed coats.

[0169] Also included within the scope of the present invention are the progeny of such plants, plant parts, and plant cells.

[0170] The transgenic plant or plant cell expressing the polypeptide or domain may be constructed in accordance with methods known in the art. In short, the plant or plant cell is constructed by incorporating one or more expression constructs encoding the polypeptide or domain into the plant host genome or chloroplast genome and propagating the resulting modified plant or plant cell into a transgenic plant or plant cell.

[0171] The expression construct is conveniently a nucleic acid construct that comprises a polynucleotide encoding a polypeptide or domain operably linked with appropriate regulatory sequences required for expression of the polynucleotide in the plant or plant part of choice. Furthermore, the expression construct may comprise a selectable marker useful for identifying plant cells into which the expression construct has been integrated and DNA sequences necessary for introduction of the construct into the plant in question (the latter depends on the DNA introduction method to be used).

[0172] The choice of regulatory sequences, such as promoter and terminator sequences and optionally signal or transit sequences, is determined, for example, on the basis of when, where, and how the polypeptide or domain is desired to be expressed. For instance, the expression of the gene encoding a polypeptide or domain may be constitutive or inducible, or may be developmental, stage or tissue specific, and the gene product may be targeted to a specific tissue or plant part such as seeds or leaves. Regulatory sequences are, for example, described by Tague et al., 1988, *Plant Physiology* 86: 506.

[0173] For constitutive expression, the 35S-CaMV, the maize ubiquitin 1, or the rice actin 1 promoter may be used (Franck et al., 1980, *Cell* 21: 285-294; Christensen et al., 1992, *Plant Mol. Biol.* 18: 675-689; Zhang et al., 1991, *Plant Cell* 3: 1155-1165). Organ-specific promoters may be, for example, a promoter from storage sink tissues such as seeds, potato tubers, and fruits (Edwards and Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303), or from metabolic sink tissues such as meristems (Ito et al., 1994, *Plant Mol. Biol.* 24: 863-878), a seed specific promoter such as the glutelin, prolamin, globulin, or albumin promoter from rice (Wu et al., 1998, *Plant Cell Physiol.* 39: 885-889), a *Vicia faba* promoter from the legumin B4 and the unknown seed protein gene from *Vicia faba* (Conrad et al., 1998, *J. Plant Physiol.* 152: 708-711), a promoter from a seed oil body protein (Chen et al., 1998, *Plant Cell Physiol.* 39: 935-941), the storage protein *napA* promoter from *Brassica napus*, or any other seed specific promoter known in the art, e.g., as described in WO 91/14772. Furthermore, the promoter may be a leaf specific promoter such as the *rbcs* promoter from rice or tomato (Kyoizuka et al., 1993, *Plant Physiol.* 102: 991-1000), the chlorella virus adenine methyltransferase gene promoter (Mitra and Higgins, 1994, *Plant Mol. Biol.* 26: 85-93), the *aldP* gene promoter from rice (Kagaya et al., 1995, *Mol. Gen. Genet.* 248: 668-674), or a wound inducible promoter such as the potato *pin2* promoter (Xu et al., 1993, *Plant Mol. Biol.* 22: 573-588). Likewise, the promoter may be induced by abiotic treatments such as temperature, drought, or alterations in salinity or induced by exogenously applied substances that activate the promoter, e.g., ethanol,

oestrogens, plant hormones such as ethylene, abscisic acid, and gibberellic acid, and heavy metals.

[0174] A promoter enhancer element may also be used to achieve higher expression of a polypeptide or domain in the plant. For instance, the promoter enhancer element may be an intron that is placed between the promoter and the polynucleotide encoding a polypeptide or domain. For instance, Xu *et al.*, 1993, *supra*, disclose the use of the first intron of the rice actin 1 gene to enhance expression.

[0175] The selectable marker gene and any other parts of the expression construct may be chosen from those available in the art.

[0176] The nucleic acid construct is incorporated into the plant genome according to conventional techniques known in the art, including *Agrobacterium*-mediated transformation, virus-mediated transformation, microinjection, particle bombardment, biolistic transformation, and electroporation (Gasser *et al.*, 1990, *Science* 244: 1293; Potrykus, 1990, *Bio/Technology* 8: 535; Shimamoto *et al.*, 1989, *Nature* 338: 274).

[0177] *Agrobacterium tumefaciens*-mediated gene transfer is a method for generating transgenic dicots (for a review, see Hooykas and Schilperoort, 1992, *Plant Mol. Biol.* 19: 15-38) and for transforming monocots, although other transformation methods may be used for these plants. A method for generating transgenic monocots is particle bombardment (microscopic gold or tungsten particles coated with the transforming DNA) of embryonic calli or developing embryos (Christou, 1992, *Plant J.* 2: 275-281; Shimamoto, 1994, *Curr. Opin. Biotechnol.* 5: 158-162; Vasil *et al.*, 1992, *Bio/Technology* 10: 667-674). An alternative method for transformation of monocots is based on protoplast transformation as described by Omirulleh *et al.*, 1993, *Plant Mol. Biol.* 21: 415-428. Additional transformation methods include those described in U.S. Patent Nos. 6,395,966 and 7,151,204.

[0178] Following transformation, the transformants having incorporated the expression construct are selected and regenerated into whole plants according to methods well known in the art. Often the transformation procedure is designed for the selective elimination of selection genes either during regeneration or in the following generations by using, for example, co-transformation with two separate T-DNA constructs or site specific excision of the selection gene by a specific recombinase.

[0179] In addition to direct transformation of a particular plant genotype with a construct of the present invention, transgenic plants may be made by crossing a plant having the construct to a second plant lacking the construct. For example, a construct encoding a polypeptide or domain can be introduced into a particular plant variety by crossing, without the need for ever directly transforming a plant of that given variety. Therefore, the present invention encompasses not only a plant directly regenerated from cells which have been transformed in accordance with the present invention, but also the progeny of such plants. As used herein, progeny may refer to the offspring of any generation of a parent plant prepared in accordance with the present

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

[0180] Plants may be generated through a process of backcross conversion. For example, plants include plants referred to as a backcross converted genotype, line, inbred, or hybrid.

[0181] Genetic markers may be used to assist in the introgression of one or more transgenes of the invention from one genetic background into another. Marker assisted selection offers advantages relative to conventional breeding in that it can be used to avoid errors caused by phenotypic variations. Further, genetic markers may provide data regarding the relative degree of elite germplasm in the individual progeny of a particular cross. For example, when a plant with a desired trait which otherwise has a non-agronomically desirable genetic background is crossed to an elite parent, genetic markers may be used to select progeny which not only possess the trait of interest, but also have a relatively large proportion of the desired germplasm. In this way, the number of generations required to introgress one or more traits into a particular genetic background is minimized.

[0182] The present disclosure also relates to methods of producing a polypeptide or domain of the present invention comprising: (a) cultivating a transgenic plant or a plant cell comprising a polynucleotide encoding the polypeptide or domain under conditions conducive for production of the polypeptide or domain; and (b) recovering the polypeptide or domain.

Fermentation Broth Formulations or Cell Compositions

[0183] The present disclosure also relates to a fermentation broth formulation or a cell composition comprising a polypeptide of the present invention. The fermentation broth product further comprises additional ingredients used in the fermentation process, such as, for example, cells (including, the host cells containing the gene encoding the polypeptide of the present invention which are used to produce the polypeptide of interest), cell debris, biomass, fermentation media and/or fermentation products. In some embodiments, the composition is a cell-killed whole broth containing organic acid(s), killed cells and/or cell debris, and culture medium.

[0184] The term "fermentation broth" as used herein refers to a preparation produced by cellular fermentation that undergoes no or minimal recovery and/or purification. For example, fermentation broths are produced when microbial cultures are grown to saturation, incubated under carbon-limiting conditions to allow protein synthesis (e.g., expression of enzymes by host cells) and secretion into cell culture medium. The fermentation broth can contain unfractionated or fractionated contents of the fermentation materials derived at the end of the fermentation. Typically, the fermentation broth is unfractionated and comprises the spent culture medium and cell debris present after the microbial cells (e.g., filamentous fungal cells)

are removed, e.g., by centrifugation. In some aspects, the fermentation broth contains spent cell culture medium, extracellular enzymes, and viable and/or nonviable microbial cells.

[0185] In an aspect, the fermentation broth formulation and cell compositions comprise a first organic acid component comprising at least one 1-5 carbon organic acid and/or a salt thereof and a second organic acid component comprising at least one 6 or more carbon organic acid and/or a salt thereof. In a specific aspect, the first organic acid component is acetic acid, formic acid, propionic acid, a salt thereof, or a mixture of two or more of the foregoing and the second organic acid component is benzoic acid, cyclohexanecarboxylic acid, 4-methylvaleric acid, phenylacetic acid, a salt thereof, or a mixture of two or more of the foregoing.

[0186] In one aspect, the composition contains an organic acid(s), and optionally further contains killed cells and/or cell debris. In one aspect, the killed cells and/or cell debris are removed from a cell-killed whole broth to provide a composition that is free of these components.

[0187] The fermentation broth formulations or cell compositions may further comprise a preservative and/or anti-microbial (e.g., bacteriostatic) agent, including, but not limited to, sorbitol, sodium chloride, potassium sorbate, and others known in the art.

[0188] The cell-killed whole broth or composition may contain the unfractionated contents of the fermentation materials derived at the end of the fermentation. Typically, the cell-killed whole broth or composition contains the spent culture medium and cell debris present after the microbial cells (e.g., filamentous fungal cells) are grown to saturation, incubated under carbon-limiting conditions to allow protein synthesis. In some aspects, the cell-killed whole broth or composition contains the spent cell culture medium, extracellular enzymes, and killed filamentous fungal cells. In some aspects, the microbial cells present in the cell-killed whole broth or composition can be permeabilized and/or lysed using methods known in the art.

[0189] A whole broth or cell composition as described herein is typically a liquid, but may contain insoluble components, such as killed cells, cell debris, culture media components, and/or insoluble enzyme(s). In some aspects, insoluble components may be removed to provide a clarified liquid composition.

[0190] The whole broth formulations and cell compositions of the present invention may be produced by a method described in WO 90/15861 or WO 2010/096673.

Enzyme Compositions

[0191] The present invention also relates to compositions comprising a polypeptide of the present invention. Preferably, the compositions are enriched in such a polypeptide. The term "enriched" indicates that the endoglucanase activity of the composition has been increased, e.g., with an enrichment factor of at least 1.1.

[0192] The compositions may comprise a polypeptide of the present invention as the major enzymatic component, e.g., a mono-component composition. Alternatively, the compositions may comprise multiple enzymatic activities, such as one or more (e.g., several) enzymes selected from the group consisting of hydrolase, isomerase, ligase, lyase, oxidoreductase, or transferase, e.g., an alpha-galactosidase, alpha-glucosidase, aminopeptidase, amylase, beta-galactosidase, beta-glucosidase, beta-xylosidase, carbohydrase, carboxypeptidase, catalase, cellobiohydrolase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, endoglucanase, esterase, glucoamylase, invertase, laccase, lipase, mannosidase, mutanase, oxidase, pectinolytic enzyme, peroxidase, phytase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transglutaminase, or xylanase.

[0193] The compositions may be prepared in accordance with methods known in the art and may be in the form of a liquid or a dry composition. The compositions may be stabilized in accordance with methods known in the art.

[0194] Examples are given below of preferred uses of the compositions of the present invention. The dosage of the composition and other conditions under which the composition is used may be determined on the basis of methods known in the art.

Uses

[0195] The present invention is also directed to the following methods for using the polypeptides having endoglucanase activity, or compositions thereof.

Processing of Cellulosic Material

[0196] The present invention also relates to methods for degrading or converting a cellulosic material, comprising: treating the cellulosic material with an enzyme composition of the present invention. In one aspect, the methods further comprise recovering the degraded or converted cellulosic material. Soluble products of degradation or conversion of the cellulosic material can be separated from insoluble cellulosic material using a method known in the art such as, for example, centrifugation, filtration, or gravity settling.

[0197] The present invention also relates to methods of producing a fermentation product, comprising: (a) saccharifying a cellulosic material with an enzyme composition of the present invention; (b) fermenting the saccharified cellulosic material with one or more (e.g., several) fermenting microorganisms to produce the fermentation product; and (c) recovering the fermentation product from the fermentation.

[0198] The present invention also relates to methods of fermenting a cellulosic material,

comprising: fermenting the cellulosic material with one or more (e.g., several) fermenting microorganisms, wherein the cellulosic material is saccharified with an enzyme composition of the present invention. In one aspect, the fermenting of the cellulosic material produces a fermentation product. In another aspect, the methods further comprise recovering the fermentation product from the fermentation.

[0199] The methods of the present invention can be used to saccharify the cellulosic material to fermentable sugars and to convert the fermentable sugars to many useful fermentation products, e.g., fuel, potable ethanol, and/or platform chemicals (e.g., acids, alcohols, ketones, gases, and the like). The production of a desired fermentation product from the cellulosic material typically involves pretreatment, enzymatic hydrolysis (saccharification), and fermentation.

[0200] The processing of the cellulosic material according to the present invention can be accomplished using processes conventional in the art. Moreover, the methods of the present invention can be implemented using any conventional biomass processing apparatus configured to operate in accordance with the invention.

[0201] Hydrolysis (saccharification) and fermentation, separate or simultaneous, include, but are not limited to, separate hydrolysis and fermentation (SHF); simultaneous saccharification and fermentation (SSF); simultaneous saccharification and co-fermentation (SSCF); hybrid hydrolysis and fermentation (HHF); separate hydrolysis and co-fermentation (SHCF); hybrid hydrolysis and co-fermentation (HHCF); and direct microbial conversion (DMC), also sometimes called consolidated bioprocessing (CBP). SHF uses separate process steps to first enzymatically hydrolyze the cellulosic material to fermentable sugars, e.g., glucose, cellobiose, and pentose monomers, and then ferment the fermentable sugars to ethanol. In SSF, the enzymatic hydrolysis of the cellulosic material and the fermentation of sugars to ethanol are combined in one step (Philippidis, G. P., 1996, Cellulose bioconversion technology, in Handbook on Bioethanol: Production and Utilization, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212). SSCF involves the co-fermentation of multiple sugars (Sheehan, J., and Himmel, M., 1999, Enzymes, energy and the environment: A strategic perspective on the U.S. Department of Energy's research and development activities for bioethanol, Biotechnol. Prog. 15: 817-827). HHF involves a separate hydrolysis step, and in addition a simultaneous saccharification and hydrolysis step, which can be carried out in the same reactor. The steps in an HHF process can be carried out at different temperatures, i.e., high temperature enzymatic saccharification followed by SSF at a lower temperature that the fermentation strain can tolerate. DMC combines all three processes (enzyme production, hydrolysis, and fermentation) in one or more (e.g., several) steps where the same organism is used to produce the enzymes for conversion of the cellulosic material to fermentable sugars and to convert the fermentable sugars into a final product (Lynd, L. R., Weimer, P. J., van Zyl, W. H., and Pretorius, I. S., 2002, Microbial cellulose utilization: Fundamentals and biotechnology, Microbiol. Mol. Biol. Reviews 66: 506-577). It is understood herein that any method known in the art comprising pretreatment, enzymatic hydrolysis (saccharification), fermentation, or a combination thereof, can be used in the practicing the methods of the

present invention.

[0202] A conventional apparatus can include a fed-batch stirred reactor, a batch stirred reactor, a continuous flow stirred reactor with ultrafiltration, and/or a continuous plug-flow column reactor (Fernanda de Castilhos Corazza, Flávio Faria de Moraes, Gisella Maria Zanin and Ivo Neitzel, 2003, Optimal control in fed-batch reactor for the cellobiose hydrolysis, *Acta Scientiarum. Technology* 25: 33-38; Gusakov, A. V., and Sinitsyn, A. P., 1985, Kinetics of the enzymatic hydrolysis of cellulose: 1. A mathematical model for a batch reactor process, *Enz. Microb. Technol.* 7: 346-352), an attrition reactor (Ryu, S. K., and Lee, J. M., 1983, Bioconversion of waste cellulose by using an attrition bioreactor, *Biotechnol. Bioeng.* 25: 53-65), or a reactor with intensive stirring induced by an electromagnetic field (Gusakov, A. V., Sinitsyn, A. P., Davydkin, I. Y., Davydkin, V. Y., Protas, O. V., 1996, Enhancement of enzymatic cellulose hydrolysis using a novel type of bioreactor with intensive stirring induced by electromagnetic field, *Appl. Biochem. Biotechnol.* 56: 141-153). Additional reactor types include: fluidized bed, upflow blanket, immobilized, and extruder type reactors for hydrolysis and/or fermentation.

[0203] Pretreatment. In practicing the methods of the present invention, any pretreatment process known in the art can be used to disrupt plant cell wall components of the cellulosic material (Chandra et al., 2007, Substrate pretreatment: The key to effective enzymatic hydrolysis of lignocellulosics? *Adv. Biochem. Engin./Biotechnol.* 108: 67-93; Galbe and Zacchi, 2007, Pretreatment of lignocellulosic materials for efficient bioethanol production, *Adv. Biochem. Engin./Biotechnol.* 108: 41-65; Hendriks and Zeeman, 2009, Pretreatments to enhance the digestibility of lignocellulosic biomass, *Bioresource Technol.* 100: 10-18; Mosier et al., 2005, Features of promising technologies for pretreatment of lignocellulosic biomass, *Bioresource Technol.* 96: 673-686; Taherzadeh and Karimi, 2008, Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: A review, *Int. J. of Mol. Sci.* 9: 1621-1651; Yang and Wyman, 2008, Pretreatment: the key to unlocking low-cost cellulosic ethanol, *Biofuels Bioproducts and Biorefining-Biofpr.* 2: 26-40).

[0204] The cellulosic material can also be subjected to particle size reduction, sieving, presoaking, wetting, washing, and/or conditioning prior to pretreatment using methods known in the art.

[0205] Conventional pretreatments include, but are not limited to, steam pretreatment (with or without explosion), dilute acid pretreatment, hot water pretreatment, alkaline pretreatment, lime pretreatment, wet oxidation, wet explosion, ammonia fiber explosion, organosolv pretreatment, and biological pretreatment. Additional pretreatments include ammonia percolation, ultrasound, electroporation, microwave, supercritical CO₂, supercritical H₂O, ozone, ionic liquid, and gamma irradiation pretreatments.

[0206] The cellulosic material can be pretreated before hydrolysis and/or fermentation. Pretreatment is preferably performed prior to the hydrolysis. Alternatively, the pretreatment can be carried out simultaneously with enzyme hydrolysis to release fermentable sugars, such as

glucose, xylose, and/or cellobiose. In most cases the pretreatment step itself results in some conversion of biomass to fermentable sugars (even in absence of enzymes).

[0207] Steam Pretreatment. In steam pretreatment, the cellulosic material is heated to disrupt the plant cell wall components, including lignin, hemicellulose, and cellulose to make the cellulose and other fractions, e.g., hemicellulose, accessible to enzymes. The cellulosic material is passed to or through a reaction vessel where steam is injected to increase the temperature to the required temperature and pressure and is retained therein for the desired reaction time. Steam pretreatment is preferably performed at 140-250°C, e.g., 160-200°C or 170-190°C, where the optimal temperature range depends on addition of a chemical catalyst. Residence time for the steam pretreatment is preferably 1-60 minutes, e.g., 1-30 minutes, 1-20 minutes, 3-12 minutes, or 4-10 minutes, where the optimal residence time depends on temperature range and addition of a chemical catalyst. Steam pretreatment allows for relatively high solids loadings, so that the cellulosic material is generally only moist during the pretreatment. The steam pretreatment is often combined with an explosive discharge of the material after the pretreatment, which is known as steam explosion, that is, rapid flashing to atmospheric pressure and turbulent flow of the material to increase the accessible surface area by fragmentation (Duff and Murray, 1996, *Bioresource Technology* 855: 1-33; Galbe and Zacchi, 2002, *Appl. Microbiol. Biotechnol.* 59: 618-628; U.S. Patent Application No. 20020164730). During steam pretreatment, hemicellulose acetyl groups are cleaved and the resulting acid autocatalyzes partial hydrolysis of the hemicellulose to monosaccharides and oligosaccharides. Lignin is removed to only a limited extent.

[0208] Chemical Pretreatment: The term "chemical treatment" refers to any chemical pretreatment that promotes the separation and/or release of cellulose, hemicellulose, and/or lignin. Such a pretreatment can convert crystalline cellulose to amorphous cellulose. Examples of suitable chemical pretreatment processes include, for example, dilute acid pretreatment, lime pretreatment, wet oxidation, ammonia fiber/freeze explosion (AFEX), ammonia percolation (APR), ionic liquid, and organosolv pretreatments.

[0209] A catalyst such as H₂SO₄ or SO₂ (typically 0.3 to 5% w/w) is often added prior to steam pretreatment, which decreases the time and temperature, increases the recovery, and improves enzymatic hydrolysis (Ballesteros et al., 2006, *Appl. Biochem. Biotechnol.* 129-132: 496-508; Varga et al., 2004, *Appl. Biochem. Biotechnol.* 113-116: 509-523; Sassner et al., 2006, *Enzyme Microb. Technol.* 39: 756-762). In dilute acid pretreatment, the cellulosic material is mixed with dilute acid, typically H₂SO₄, and water to form a slurry, heated by steam to the desired temperature, and after a residence time flashed to atmospheric pressure. The dilute acid pretreatment can be performed with a number of reactor designs, e.g., plug-flow reactors, counter-current reactors, or continuous counter-current shrinking bed reactors (Duff and Murray, 1996, *supra*; Schell et al., 2004, *Bioresource Technol.* 91: 179-188; Lee et al., 1999, *Adv. Biochem. Eng. Biotechnol.* 65: 93-115).

[0210] Several methods of pretreatment under alkaline conditions can also be used. These alkaline pretreatments include, but are not limited to, sodium hydroxide, lime, wet oxidation,

ammonia percolation (APR), and ammonia fiber/freeze explosion (AFEX).

[0211] Lime pretreatment is performed with calcium oxide or calcium hydroxide at temperatures of 85-150°C and residence times from 1 hour to several days (Wyman et al., 2005, *Bioresource Technol.* 96: 1959-1966; Mosier et al., 2005, *Bioresource Technol.* 96: 673-686). WO 2006/110891, WO 2006/110899, WO 2006/110900, and WO 2006/110901 disclose pretreatment methods using ammonia.

[0212] Wet oxidation is a thermal pretreatment performed typically at 180-200°C for 5-15 minutes with addition of an oxidative agent such as hydrogen peroxide or over-pressure of oxygen (Schmidt and Thomsen, 1998, *Bioresource Technol.* 64: 139-151; Palonen et al., 2004, *Appl. Biochem. Biotechnol.* 117: 1-17; Varga et al., 2004, *Biotechnol. Bioeng.* 88: 567-574; Martin et al., 2006, *J. Chem. Technol. Biotechnol.* 81: 1669-1677). The pretreatment is performed preferably at 1-40% dry matter, e.g., 2-30% dry matter or 5-20% dry matter, and often the initial pH is increased by the addition of alkali such as sodium carbonate.

[0213] A modification of the wet oxidation pretreatment method, known as wet explosion (combination of wet oxidation and steam explosion), can handle dry matter up to 30%. In wet explosion, the oxidizing agent is introduced during pretreatment after a certain residence time. The pretreatment is then ended by flashing to atmospheric pressure (WO 2006/032282).

[0214] Ammonia fiber explosion (AFEX) involves treating the cellulosic material with liquid or gaseous ammonia at moderate temperatures such as 90-150°C and high pressure such as 17-20 bar for 5-10 minutes, where the dry matter content can be as high as 60% (Gollapalli et al., 2002, *Appl. Biochem. Biotechnol.* 98: 23-35; Chundawat et al., 2007, *Biotechnol. Bioeng.* 96: 219-231; Alizadeh et al., 2005, *Appl. Biochem. Biotechnol.* 121: 1133-1141; Teymouri et al., 2005, *Bioresource Technol.* 96: 2014-2018). During AFEX pretreatment cellulose and hemicelluloses remain relatively intact. Lignin-carbohydrate complexes are cleaved.

[0215] Organosolv pretreatment delignifies the cellulosic material by extraction using aqueous ethanol (40-60% ethanol) at 160-200°C for 30-60 minutes (Pan et al., 2005, *Biotechnol. Bioeng.* 90: 473-481; Pan et al., 2006, *Biotechnol. Bioeng.* 94: 851-861; Kurabi et al., 2005, *Appl. Biochem. Biotechnol.* 121: 219-230). Sulphuric acid is usually added as a catalyst. In organosolv pretreatment, the majority of hemicellulose and lignin is removed.

[0216] Other examples of suitable pretreatment methods are described by Schell et al., 2003, *Appl. Biochem. and Biotechnol.* Vol. 105-108, p. 69-85, and Mosier et al., 2005, *Bioresource Technology* 96: 673-686, and U.S. Published Application 2002/0164730.

[0217] In one aspect, the chemical pretreatment is preferably carried out as a dilute acid treatment, and more preferably as a continuous dilute acid treatment. The acid is typically sulfuric acid, but other acids can also be used, such as acetic acid, citric acid, nitric acid, phosphoric acid, tartaric acid, succinic acid, hydrogen chloride, or mixtures thereof. Mild acid treatment is conducted in the pH range of preferably 1-5, e.g., 1-4 or 1-2.5. In one aspect, the

acid concentration is in the range from preferably 0.01 to 10 wt % acid, e.g., 0.05 to 5 wt % acid or 0.1 to 2 wt % acid. The acid is contacted with the cellulosic material and held at a temperature in the range of preferably 140-200°C, e.g., 165-190°C, for periods ranging from 1 to 60 minutes.

[0218] In another aspect, pretreatment takes place in an aqueous slurry. In preferred aspects, the cellulosic material is present during pretreatment in amounts preferably between 10-80 wt %, e.g., 20-70 wt % or 30-60 wt %, such as around 40 wt %. The pretreated cellulosic material can be unwashed or washed using any method known in the art, e.g., washed with water.

[0219] Mechanical Pretreatment or Physical Pretreatment: The term "mechanical pretreatment" or "physical pretreatment" refers to any pretreatment that promotes size reduction of particles. For example, such pretreatment can involve various types of grinding or milling (e.g., dry milling, wet milling, or vibratory ball milling).

[0220] The cellulosic material can be pretreated both physically (mechanically) and chemically. Mechanical or physical pretreatment can be coupled with steaming/steam explosion, hydrothermolysis, dilute or mild acid treatment, high temperature, high pressure treatment, irradiation (e.g., microwave irradiation), or combinations thereof. In one aspect, high pressure means pressure in the range of preferably about 100 to about 400 psi, e.g., about 150 to about 250 psi. In another aspect, high temperature means temperatures in the range of about 100 to about 300°C, e.g., about 140 to about 200°C. In a preferred aspect, mechanical or physical pretreatment is performed in a batch-process using a steam gun hydrolyzer system that uses high pressure and high temperature as defined above, e.g., a Sunds Hydrolyzer available from Sunds Defibrator AB, Sweden. The physical and chemical pretreatments can be carried out sequentially or simultaneously, as desired.

[0221] Accordingly, in a preferred aspect, the cellulosic material is subjected to physical (mechanical) or chemical pretreatment, or any combination thereof, to promote the separation and/or release of cellulose, hemicellulose, and/or lignin.

[0222] Biological Pretreatment: The term "biological pretreatment" refers to any biological pretreatment that promotes the separation and/or release of cellulose, hemicellulose, and/or lignin from the cellulosic material. Biological pretreatment techniques can involve applying lignin-solubilizing microorganisms and/or enzymes (see, for example, Hsu, T.-A., 1996, Pretreatment of biomass, in Handbook on Bioethanol: Production and Utilization, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212; Ghosh and Singh, 1993, Physicochemical and biological treatments for enzymatic/microbial conversion of cellulosic biomass, Adv. Appl. Microbiol. 39: 295-333; McMillan, J. D., 1994, Pretreating lignocellulosic biomass: a review, in Enzymatic Conversion of Biomass for Fuels Production, Himmel, M. E., Baker, J. O., and Overend, R. P., eds., ACS Symposium Series 566, American Chemical Society, Washington, DC, chapter 15; Gong, C. S., Cao, N. J., Du, J., and Tsao, G. T., 1999, Ethanol production from renewable resources, in Advances in Biochemical Engineering/Biotechnology, Scheper, T., ed., Springer-Verlag Berlin Heidelberg, Germany, 65: 207-241; Olsson and Hahn-Hagerdal, 1996,

Fermentation of lignocellulosic hydrolysates for ethanol production, *Enz. Microb. Tech.* 18: 312-331; and Vallander and Eriksson, 1990, Production of ethanol from lignocellulosic materials: State of the art, *Adv. Biochem. Eng./Biotechnol.* 42: 63-95).

[0223] Saccharification. In the hydrolysis step, also known as saccharification, the cellulosic material, *e.g.*, pretreated, is hydrolyzed to break down cellulose and/or hemicellulose to fermentable sugars, such as glucose, cellobiose, xylose, xylulose, arabinose, mannose, galactose, and/or soluble oligosaccharides. The hydrolysis is performed enzymatically by an enzyme composition in the presence of a polypeptide having endoglucanase activity of the present invention. The enzymes of the compositions can be added simultaneously or sequentially.

[0224] Enzymatic hydrolysis is preferably carried out in a suitable aqueous environment under conditions that can be readily determined by one skilled in the art. In one aspect, hydrolysis is performed under conditions suitable for the activity of the enzyme(s), *i.e.*, optimal for the enzyme(s). The hydrolysis can be carried out as a fed batch or continuous process where the cellulosic material is fed gradually to, for example, an enzyme containing hydrolysis solution.

[0225] The saccharification is generally performed in stirred-tank reactors or fermentors under controlled pH, temperature, and mixing conditions. Suitable process time, temperature and pH conditions can readily be determined by one skilled in the art. For example, the saccharification can last up to 200 hours, but is typically performed for preferably about 12 to about 120 hours, *e.g.*, about 16 to about 72 hours or about 24 to about 48 hours. The temperature is in the range of preferably about 25°C to about 70°C, *e.g.*, about 30°C to about 65°C, about 40°C to about 60°C, or about 50°C to about 55°C. The pH is in the range of preferably about 3 to about 8, *e.g.*, about 3.5 to about 7, about 4 to about 6, or about 5.0 to about 5.5. The dry solids content is in the range of preferably about 5 to about 50 wt %, *e.g.*, about 10 to about 40 wt % or about 20 to about 30 wt %.

[0226] The enzyme compositions can comprise any protein useful in degrading or converting the cellulosic material.

[0227] In one aspect, the enzyme composition comprises or further comprises one or more (*e.g.*, several) proteins selected from the group consisting of a cellulase, a GH61 polypeptide having cellulolytic enhancing activity, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin. In another aspect, the cellulase is preferably one or more (*e.g.*, several) enzymes selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase. In another aspect, the hemicellulase is preferably one or more (*e.g.*, several) enzymes selected from the group consisting of an acetylmannan esterase, an acetylxylan esterase, an arabinanase, an arabinofuranosidase, a coumaric acid esterase, a feruloyl esterase, a galactosidase, a glucuronidase, a glucuronoyl esterase, a mannanase, a mannosidase, a xylanase, and a xylosidase.

[0228] In another aspect, the enzyme composition comprises one or more (e.g., several) cellulolytic enzymes. In another aspect, the enzyme composition comprises or further comprises one or more (e.g., several) hemicellulolytic enzymes. In another aspect, the enzyme composition comprises one or more (e.g., several) cellulolytic enzymes and one or more (e.g., several) hemicellulolytic enzymes. In another aspect, the enzyme composition comprises one or more (e.g., several) enzymes selected from the group of cellulolytic enzymes and hemicellulolytic enzymes. In another aspect, the enzyme composition comprises an endoglucanase. In another aspect, the enzyme composition comprises a cellobiohydrolase. In another aspect, the enzyme composition comprises a beta-glucosidase. In another aspect, the enzyme composition comprises a polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises an endoglucanase and a polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises a cellobiohydrolase and a polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises a beta-glucosidase and a polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises an endoglucanase and a cellobiohydrolase. In another aspect, the enzyme composition comprises an endoglucanase and a beta-glucosidase. In another aspect, the enzyme composition comprises a cellobiohydrolase and a beta-glucosidase. In another aspect, the enzyme composition comprises an endoglucanase, a cellobiohydrolase, and a polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises an endoglucanase, a beta-glucosidase, and a polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises a cellobiohydrolase, a beta-glucosidase, and a polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises an endoglucanase, a cellobiohydrolase, and a beta-glucosidase. In another aspect, the enzyme composition comprises an endoglucanase, a cellobiohydrolase, a beta-glucosidase, and a polypeptide having cellulolytic enhancing activity.

[0229] In another aspect, the enzyme composition comprises an acetylmannan esterase. In another aspect, the enzyme composition comprises an acetylxylan esterase. In another aspect, the enzyme composition comprises an arabinanase (e.g., alpha-L-arabinanase). In another aspect, the enzyme composition comprises an arabinofuranosidase (e.g., alpha-L-arabinofuranosidase). In another aspect, the enzyme composition comprises a coumaric acid esterase. In another aspect, the enzyme composition comprises a feruloyl esterase. In another aspect, the enzyme composition comprises a galactosidase (e.g., alpha-galactosidase and/or beta-galactosidase). In another aspect, the enzyme composition comprises a glucuronidase (e.g., alpha-D-glucuronidase). In another aspect, the enzyme composition comprises a glucuronoyl esterase. In another aspect, the enzyme composition comprises a mannanase. In another aspect, the enzyme composition comprises a mannosidase (e.g., beta-mannosidase). In another aspect, the enzyme composition comprises a xylanase. In a preferred aspect, the xylanase is a Family 10 xylanase. In another aspect, the enzyme composition comprises a xylosidase (e.g., beta-xylosidase).

[0230] In another aspect, the enzyme composition comprises an esterase. In another aspect, the enzyme composition comprises an expansin. In another aspect, the enzyme composition

comprises a laccase. In another aspect, the enzyme composition comprises a ligninolytic enzyme. In a preferred aspect, the ligninolytic enzyme is a manganese peroxidase. In another preferred aspect, the ligninolytic enzyme is a lignin peroxidase. In another preferred aspect, the ligninolytic enzyme is a H₂O₂-producing enzyme. In another aspect, the enzyme composition comprises a pectinase. In another aspect, the enzyme composition comprises a peroxidase. In another aspect, the enzyme composition comprises a protease. In another aspect, the enzyme composition comprises a swollenin.

[0231] In the methods of the present invention, the enzyme(s) can be added prior to or during saccharification, saccharification and fermentation, or fermentation.

[0232] One or more (e.g., several) components of the enzyme composition may be wild-type proteins, recombinant proteins, or a combination of wild-type proteins and recombinant proteins. For example, one or more (e.g., several) components may be native proteins of a cell, which is used as a host cell to express recombinantly one or more (e.g., several) other components of the enzyme composition. One or more (e.g., several) components of the enzyme composition may be produced as monocomponents, which are then combined to form the enzyme composition. The enzyme composition may be a combination of multicomponent and monocomponent protein preparations.

[0233] The enzymes used in the methods of the present invention may be in any form suitable for use, such as, for example, a fermentation broth formulation or a cell composition, a cell lysate with or without cellular debris, a semi-purified or purified enzyme preparation, or a host cell as a source of the enzymes. The enzyme composition may be a dry powder or granulate, a non-dusting granulate, a liquid, a stabilized liquid, or a stabilized protected enzyme. Liquid enzyme preparations may, for instance, be stabilized by adding stabilizers such as a sugar, a sugar alcohol or another polyol, and/or lactic acid or another organic acid according to established processes.

[0234] The optimum amounts of the enzymes and polypeptides having endoglucanase activity depend on several factors including, but not limited to, the mixture of component cellulolytic enzymes and/or hemicellulolytic enzymes, the cellulosic material, the concentration of cellulosic material, the pretreatment(s) of the cellulosic material, temperature, time, pH, and inclusion of fermenting organism (e.g., yeast for Simultaneous Saccharification and Fermentation).

[0235] In one aspect, an effective amount of cellulolytic or hemicellulolytic enzyme to the cellulosic material is about 0.5 to about 50 mg, e.g., about 0.5 to about 40 mg, about 0.5 to about 25 mg, about 0.75 to about 20 mg, about 0.75 to about 15 mg, about 0.5 to about 10 mg, or about 2.5 to about 10 mg per g of the cellulosic material.

[0236] In another aspect, an effective amount of a polypeptide having endoglucanase activity to the cellulosic material is about 0.01 to about 50.0 mg, e.g., about 0.01 to about 40 mg, about 0.01 to about 30 mg, about 0.01 to about 20 mg, about 0.01 to about 10 mg, about 0.01 to about 5 mg, about 0.025 to about 1.5 mg, about 0.05 to about 1.25 mg, about 0.075 to

about 1.25 mg, about 0.1 to about 1.25 mg, about 0.15 to about 1.25 mg, or about 0.25 to about 1.0 mg per g of the cellulosic material.

[0237] In another aspect, an effective amount of a polypeptide having endoglucanase activity to cellulolytic or hemicellulolytic enzyme is about 0.005 to about 1.0 g, e.g., about 0.01 to about 1.0 g, about 0.15 to about 0.75 g, about 0.15 to about 0.5 g, about 0.1 to about 0.5 g, about 0.1 to about 0.25 g, or about 0.05 to about 0.2 g per g of cellulolytic or hemicellulolytic enzyme.

[0238] The polypeptides having cellulolytic enzyme activity or hemicellulolytic enzyme activity as well as other proteins/polypeptides useful in the degradation of the cellulosic material, e.g., GH61 polypeptides having cellulolytic enhancing activity (collectively hereinafter "polypeptides having enzyme activity") can be derived or obtained from any suitable origin, including, bacterial, fungal, yeast, plant, or mammalian origin. The term "obtained" also means herein that the enzyme may have been produced recombinantly in a host organism employing methods described herein, wherein the recombinantly produced enzyme is either native or foreign to the host organism or has a modified amino acid sequence, e.g., having one or more (e.g., several) amino acids that are deleted, inserted and/or substituted, i.e., a recombinantly produced enzyme that is a mutant and/or a fragment of a native amino acid sequence or an enzyme produced by nucleic acid shuffling processes known in the art. Encompassed within the meaning of a native enzyme are natural variants and within the meaning of a foreign enzyme are variants obtained recombinantly, such as by site-directed mutagenesis or shuffling.

[0239] A polypeptide having enzyme activity may be a bacterial polypeptide. For example, the polypeptide may be a Gram positive bacterial polypeptide such as a *Bacillus*, *Streptococcus*, *Streptomyces*, *Staphylococcus*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Clostridium*, *Geobacillus*, *Caldicellulosiruptor*, *Acidothermus*, *Thermobifidia*, or *Oceanobacillus* polypeptide having enzyme activity, or a Gram negative bacterial polypeptide such as an *E. coli*, *Pseudomonas*, *Salmonella*, *Campylobacter*, *Helicobacter*, *Flavobacterium*, *Fusobacterium*, *Ilyobacter*, *Neisseria*, or *Ureaplasma* polypeptide having enzyme activity.

[0240] In one aspect, the polypeptide is a *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, or *Bacillus thuringiensis* polypeptide having enzyme activity.

[0241] In another aspect, the polypeptide is a *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, or *Streptococcus equi* subsp. *Zooepidemicus* polypeptide having enzyme activity.

[0242] In another aspect, the polypeptide is a *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, or *Streptomyces lividans* polypeptide having enzyme activity.

[0243] The polypeptide having enzyme activity may also be a fungal polypeptide, and more preferably a yeast polypeptide such as a *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* polypeptide having enzyme activity; or more preferably a filamentous fungal polypeptide such as an *Acremonium*, *Agaricus*, *Alternaria*, *Aspergillus*, *Aureobasidium*, *Botryosphaeria*, *Ceriporiopsis*, *Chaetomidium*, *Chrysosporium*, *Claviceps*, *Cochliobolus*, *Coprinopsis*, *Coptotermes*, *Corynascus*, *Cryphonectria*, *Cryptococcus*, *Diplodia*, *Exidia*, *Filibasidium*, *Fusarium*, *Gibberella*, *Holomastigotoides*, *Humicola*, *Irpex*, *Lentinula*, *Leptosphaeria*, *Magnaporthe*, *Melanocarpus*, *Meripilus*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Piromyces*, *Poitrasia*, *Pseudoplectania*, *Pseudotrichonympha*, *Rhizomucor*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trichoderma*, *Trichophaea*, *Verticillium*, *Volvariella*, or *Xylaria* polypeptide having enzyme activity.

[0244] In one aspect, the polypeptide is a *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, or *Saccharomyces oviformis* polypeptide having enzyme activity.

[0245] In another aspect, the polypeptide is an *Acremonium cellulolyticus*, *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium tropicum*, *Chrysosporium merdarium*, *Chrysosporium inops*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium zonatum*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola grisea*, *Humicola insolens*, *Humicola lanuginosa*, *Irpex lacteus*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium funiculosum*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Thielavia achromatica*, *Thielavia albomyces*, *Thielavia albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspora*, *Thielavia ovispora*, *Thielavia peruviana*, *Thielavia spededonium*, *Thielavia setosa*, *Thielavia subthermophila*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, *Trichoderma viride*, or *Trichophaea saccata* polypeptide having enzyme activity.

[0246] Chemically modified or protein engineered mutants of polypeptides having enzyme activity may also be used.

[0247] One or more (e.g., several) components of the enzyme composition may be a recombinant component, i.e., produced by cloning of a DNA sequence encoding the single component and subsequent cell transformed with the DNA sequence and expressed in a host (see, for example, WO 91/17243 and WO 91/17244). The host is preferably a heterologous host (enzyme is foreign to host), but the host may under certain conditions also be a

homologous host (enzyme is native to host). Monocomponent cellulolytic proteins may also be prepared by purifying such a protein from a fermentation broth.

[0248] In one aspect, the one or more (e.g., several) cellulolytic enzymes comprise a commercial cellulolytic enzyme preparation. Examples of commercial cellulolytic enzyme preparations suitable for use in the present invention include, for example, CELLIC® CTec (Novozymes A/S), CELLIC® CTec2 (Novozymes A/S), CELLIC® CTec3 (Novozymes A/S), CELLUCLAST™ (Novozymes A/S), NOVOZYM™ 188 (Novozymes A/S), CELLUZYME™ (Novozymes A/S), CEREFLO™ (Novozymes A/S), and ULTRAFLO™ (Novozymes A/S), ACCELERASE™ (Genencor Int.), LAMINEX™ (Genencor Int.), SPEZYME™ CP (Genencor Int.), FILTRASE® NL (DSM); METHAPLUS® S/L 100 (DSM), ROHAMENT™ 7069 W (Rohm GmbH), FIBREZYME® LDI (Dyadic International, Inc.), FIBREZYME® LBR (Dyadic International, Inc.), or VISCOSTAR® 150L (Dyadic International, Inc.). The cellulase enzymes are added in amounts effective from about 0.001 to about 5.0 wt % of solids, e.g., about 0.025 to about 4.0 wt % of solids or about 0.005 to about 2.0 wt % of solids.

[0249] Examples of bacterial endoglucanases that can be used in the methods of the present invention, include, but are not limited to, an *Acidothermus cellulolyticus* endoglucanase (WO 91/05039; WO 93/15186; U.S. Patent No. 5,275,944; WO 96/02551; U.S. Patent No. 5,536,655, WO 00/70031, WO 05/093050); *Thermobifida fusca* endoglucanase III (WO 05/093050); and *Thermobifida fusca* endoglucanase V (WO 05/093050).

[0250] Examples of fungal endoglucanases that can be used in the present invention, include, but are not limited to, a *Trichoderma reesei* endoglucanase I (Penttilä et al., 1986, Gene 45: 253-263, *Trichoderma reesei* Cel7B endoglucanase I (GENBANK™ accession no. M15665), *Trichoderma reesei* endoglucanase II (Saloheimo, et al., 1988, Gene 63:11-22), *Trichoderma reesei* Cel5A endoglucanase II (GENBANK™ accession no. M19373), *Trichoderma reesei* endoglucanase III (Okada et al., 1988, Appl. Environ. Microbiol. 64: 555-563, GENBANK™ accession no. AB003694), *Trichoderma reesei* endoglucanase V (Saloheimo et al., 1994, Molecular Microbiology 13: 219-228, GENBANK™ accession no. Z33381), *Aspergillus aculeatus* endoglucanase (Ooi et al., 1990, Nucleic Acids Research 18: 5884), *Aspergillus kawachii* endoglucanase (Sakamoto et al., 1995, Current Genetics 27: 435-439), *Erwinia carotovora* endoglucanase (Saarilahti et al., 1990, Gene 90: 9-14), *Fusarium oxysporum* endoglucanase (GENBANK™ accession no. L29381), *Humicola grisea* var. *thermoidea* endoglucanase (GENBANK™ accession no. AB003107), *Melanocarpus albomyces* endoglucanase (GENBANK™ accession no. MAL515703), *Neurospora crassa* endoglucanase (GENBANK™ accession no. XM_324477), *Humicola insolens* endoglucanase V, *Myceliophthora thermophila* CBS 117.65 endoglucanase, basidiomycete CBS 495.95 endoglucanase, basidiomycete CBS 494.95 endoglucanase, *Thielavia terrestris* NRRL 8126 CEL6B endoglucanase, *Thielavia terrestris* NRRL 8126 CEL6C endoglucanase, *Thielavia terrestris* NRRL 8126 CEL7C endoglucanase, *Thielavia terrestris* NRRL 8126 CEL7E endoglucanase, *Thielavia terrestris* NRRL 8126 CEL7F endoglucanase, *Cladorrhinum foecundissimum* ATCC 62373 CEL7A endoglucanase, and *Trichoderma reesei* strain No. VTT-D-80133 endoglucanase (GENBANK™ accession no. M15665).

[0251] Examples of cellobiohydrolases useful in the present invention include, but are not limited to, *Aspergillus aculeatus* cellobiohydrolase II (WO 2011/059740), *Chaetomium thermophilum* cellobiohydrolase I, *Chaetomium thermophilum* cellobiohydrolase II, *Humicola insolens* cellobiohydrolase I, *Myceliophthora thermophila* cellobiohydrolase II (WO 2009/042871), *Thielavia hyrcanie* cellobiohydrolase II (WO 2010/141325), *Thielavia terrestris* cellobiohydrolase II (CEL6A, WO 2006/074435), *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, and *Trichophaea saccata* cellobiohydrolase II (WO 2010/057086).

[0252] Examples of beta-glucosidases useful in the present invention include, but are not limited to, beta-glucosidases from *Aspergillus aculeatus* (Kawaguchi et al., 1996, Gene 173: 287-288), *Aspergillus fumigatus* (WO 2005/047499), *Aspergillus niger* (Dan et al., 2000, J. Biol. Chem. 275: 4973-4980), *Aspergillus oryzae* (WO 2002/095014), *Penicillium brasilianum* IBT 20888 (WO 2007/019442 and WO 2010/088387), *Thielavia terrestris* (WO 2011/035029), and *Trichophaea saccata* (WO 2007/019442).

[0253] The beta-glucosidase may be a fusion protein. In one aspect, the beta-glucosidase is an *Aspergillus oryzae* beta-glucosidase variant BG fusion protein (WO 2008/057637) or an *Aspergillus oryzae* beta-glucosidase fusion protein (WO 2008/057637).

[0254] Other useful endoglucanases, cellobiohydrolases, and beta-glucosidases are disclosed in numerous Glycosyl Hydrolase families using the classification according to Henrissat B., 1991, A classification of glycosyl hydrolases based on amino-acid sequence similarities, Biochem. J. 280: 309-316, and Henrissat B., and Bairoch A., 1996, Updating the sequence-based classification of glycosyl hydrolases, Biochem. J. 316: 695-696.

[0255] Other cellulolytic enzymes that may be used in the present invention are described in WO 98/13465, WO 98/015619, WO 98/015633, WO 99/06574, WO 99/10481, WO 99/025847, WO 99/031255, WO 2002/101078, WO 2003/027306, WO 2003/052054, WO 2003/052055, WO 2003/052056, WO 2003/052057, WO 2003/052118, WO 2004/016760, WO 2004/043980, WO 2004/048592, WO 2005/001065, WO 2005/028636, WO 2005/093050, WO 2005/093073, WO 2006/074005, WO 2006/117432, WO 2007/071818, WO 2007/071820, WO 2008/008070, WO 2008/008793, U.S. Patent No. 5,457,046, U.S. Patent No. 5,648,263, and U.S. Patent No. 5,686,593.

[0256] In the methods of the present invention, any GH61 polypeptide having cellulolytic enhancing activity can be used.

[0257] In a first aspect, the GH61 polypeptide having cellulolytic enhancing activity comprises the following motifs:

[ILMV]-P-X(4,5)-G-X-Y-[ILMV]-X-R-X-[EQ]-X(4)-[HNQ] (SEQ ID NO: 21 or SEQ ID NO: 22) and [FW]-[TF]-K-[AIV], wherein X is any amino acid, X(4,5) is any amino acid at 4 or 5 contiguous positions, and X(4) is any amino acid at 4 contiguous positions.

[0258] The isolated polypeptide comprising the above-noted motifs may further comprise:

H-X(1,2)-G-P-X(3)-[YW]-[AILMV] (SEQ ID NO: 23 or SEQ ID NO: 24),

[EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV] (SEQ ID NO: 25), or

H-X(1,2)-G-P-X(3)-[YW]-[AILMV] (SEQ ID NO: 26 or SEQ ID NO: 27) and [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV] (SEQ ID NO: 28),

wherein X is any amino acid, X(1,2) is any amino acid at 1 position or 2 contiguous positions, X(3) is any amino acid at 3 contiguous positions, and X(2) is any amino acid at 2 contiguous positions. In the above motifs, the accepted IUPAC single letter amino acid abbreviation is employed.

[0259] In a preferred aspect, the isolated GH61 polypeptide having cellulolytic enhancing activity further comprises H-X(1,2)-G-P-X(3)-[YW]-[AILMV] (SEQ ID NO: 23 or SEQ ID NO: 24). In another preferred embodiment, the isolated GH61 polypeptide having cellulolytic enhancing activity further comprises [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV] (SEQ ID NO: 25). In another preferred aspect, the isolated GH61 polypeptide having cellulolytic enhancing activity further comprises H-X(1,2)-G-P-X(3)-[YW]-[AILMV] (SEQ ID NO: 26 or SEQ ID NO: 27) and [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV] (SEQ ID NO: 28).

[0260] In a second aspect, the GH61 polypeptide having cellulolytic enhancing activity comprises the following motif:

[ILMV]-P-X(4,5)-G-X-Y-[ILMV]-X-R-X-[EQ]-X(3)-A-[HNQ] (SEQ ID NO: 29 or SEQ ID NO: 30),

wherein x is any amino acid, x(4,5) is any amino acid at 4 or 5 contiguous positions, and x(3) is any amino acid at 3 contiguous positions. In the above motif, the accepted IUPAC single letter amino acid abbreviation is employed.

[0261] Examples of GH61 polypeptides having cellulolytic enhancing activity useful in the processes of the present invention include, but are not limited to, GH61 polypeptides from *Thielavia terrestris* (WO 2005/074647, WO 2008/148131, and WO 2011/035027), *Thermoascus aurantiacus* (WO 2005/074656 and WO 2010/065830), *Trichoderma reesei* (WO 2007/089290), *Myceliophthora thermophila* (WO 2009/085935, WO 2009/085859, WO 2009/085864, WO 2009/085868), *Aspergillus fumigatus* (WO 2010/138754), GH61 polypeptides from *Penicillium pinophilum* (WO 2011/005867), *Thermoascus* sp. (WO 2011/039319), *Penicillium* sp. (WO 2011/041397), and *Thermoascus crustaceus* (WO 2011/041504).

[0262] In one aspect, the GH61 polypeptide having cellulolytic enhancing activity is used in the presence of a soluble activating divalent metal cation according to WO 2008/151043, e.g., manganese sulfate.

[0263] In another aspect, the GH61 polypeptide having cellulolytic enhancing activity is used in the presence of a dioxy compound, a bicyclic compound, a heterocyclic compound, a nitrogen-

containing compound, a quinone compound, a sulfur-containing compound, or a liquor obtained from a pretreated cellulosic material such as pretreated corn stover (PCS).

[0264] The dioxy compound may include any suitable compound containing two or more oxygen atoms. In some aspects, the dioxy compounds contain a substituted aryl moiety as described herein. The dioxy compounds may comprise one or more (e.g., several) hydroxyl and/or hydroxyl derivatives, but also include substituted aryl moieties lacking hydroxyl and hydroxyl derivatives. Non-limiting examples of the dioxy compounds include pyrocatechol or catechol; caffeic acid; 3,4-dihydroxybenzoic acid; 4-tert-butyl-5-methoxy-1,2-benzenediol; pyrogallol; gallic acid; methyl-3,4,5-trihydroxybenzoate; 2,3,4-trihydroxybenzophenone; 2,6-dimethoxyphenol; sinapinic acid; 3,5-dihydroxybenzoic acid; 4-chloro-1,2-benzenediol; 4-nitro-1,2-benzenediol; tannic acid; ethyl gallate; methyl glycolate; dihydroxyfumaric acid; 2-butyne-1,4-diol; (croconic acid; 1,3-propanediol; tartaric acid; 2,4-pentanediol; 3-ethoxy-1,2-propanediol; 2,4,4'-trihydroxybenzophenone; cis-2-butene-1,4-diol; 3,4-dihydroxy-3-cyclobutene-1,2-dione; dihydroxyacetone; acrolein acetal; methyl-4-hydroxybenzoate; 4-hydroxybenzoic acid; and methyl-3,5-dimethoxy-4-hydroxybenzoate; or a salt or solvate thereof.

[0265] The bicyclic compound may include any suitable substituted fused ring system as described herein. The compounds may comprise one or more (e.g., several) additional rings, and are not limited to a specific number of rings unless otherwise stated. In one aspect, the bicyclic compound is a flavonoid. In another aspect, the bicyclic compound is an optionally substituted isoflavonoid. In another aspect, the bicyclic compound is an optionally substituted flavylum ion, such as an optionally substituted anthocyanidin or optionally substituted anthocyanin, or derivative thereof. Non-limiting examples of the bicyclic compounds include epicatechin; quercetin; myricetin; taxifolin; kaempferol; morin; acacetin; naringenin; isorhamnetin; apigenin; cyanidin; cyanin; kuromanin; keracyanin; or a salt or solvate thereof.

[0266] The heterocyclic compound may be any suitable compound, such as an optionally substituted aromatic or non-aromatic ring comprising a heteroatom, as described herein. In one aspect, the heterocyclic is a compound comprising an optionally substituted heterocycloalkyl moiety or an optionally substituted heteroaryl moiety. In another aspect, the optionally substituted heterocycloalkyl moiety or optionally substituted heteroaryl moiety is an optionally substituted 5-membered heterocycloalkyl or an optionally substituted 5-membered heteroaryl moiety. In another aspect, the optionally substituted heterocycloalkyl or optionally substituted heteroaryl moiety is an optionally substituted moiety selected from pyrazolyl, furanyl, imidazolyl, isoxazolyl, oxadiazolyl, oxazolyl, pyrrolyl, pyridyl, pyrimidyl, pyridazinyl, thiazolyl, triazolyl, thienyl, dihydrothieno-pyrazolyl, thianaphthenyl, carbazolyl, benzimidazolyl, benzothienyl, benzofuranyl, indolyl, quinolinyl, benzotriazolyl, benzothiazolyl, benzooxazolyl, benzimidazolyl, isoquinolinyl, isoindolyl, acridinyl, benzoisazolyl, dimethylhydantoin, pyrazinyl, tetrahydrofuranyl, pyrrolinyl, pyrrolidinyl, morpholinyl, indolyl, diazepinyl, azepinyl, thiepinyl, piperidinyl, and oxepinyl. In another aspect, the optionally substituted heterocycloalkyl moiety or optionally substituted heteroaryl moiety is an optionally substituted furanyl. Non-limiting examples of the heterocyclic compounds include (1,2-dihydroxyethyl)-3,4-dihydroxyfuran-

2(5H)-one; 4-hydroxy-5-methyl-3-furanone; 5-hydroxy-2(5H)-furanone; [1,2-dihydroxyethyl]furan-2,3,4(5H)-trione; α -hydroxy- γ -butyrolactone; ribonic γ -lactone; aldohexuronic acid γ -lactone; gluconic acid δ -lactone; 4-hydroxycoumarin; dihydrobenzofuran; 5-(hydroxymethyl)furfural; furoin; 2(5H)-furanone; 5,6-dihydro-2H-pyran-2-one; and 5,6-dihydro-4-hydroxy-6-methyl-2H-pyran-2-one; or a salt or solvate thereof.

[0267] The nitrogen-containing compound may be any suitable compound with one or more nitrogen atoms. In one aspect, the nitrogen-containing compound comprises an amine, imine, hydroxylamine, or nitroxide moiety. Non-limiting examples of the nitrogen-containing compounds include acetone oxime; violuric acid; pyridine-2-aldoxime; 2-aminophenol; 1,2-benzenediamine; 2,2,6,6-tetramethyl-1-piperidinyloxy; 5,6,7,8-tetrahydrobiopterin; 6,7-dimethyl-5,6,7,8-tetrahydropterine; and maleamic acid; or a salt or solvate thereof.

[0268] The quinone compound may be any suitable compound comprising a quinone moiety as described herein. Non-limiting examples of the quinone compounds include 1,4-benzoquinone; 1,4-naphthoquinone; 2-hydroxy-1,4-naphthoquinone; 2,3-dimethoxy-5-methyl-1,4-benzoquinone or coenzyme Q₀; 2,3,5,6-tetramethyl-1,4-benzoquinone or duroquinone; 1,4-dihydroxyanthraquinone; 3-hydroxy-1-methyl-5,6-indolinedione or adrenochrome; 4-tert-butyl-5-methoxy-1,2-benzoquinone; pyrroloquinoline quinone; or a salt or solvate thereof.

[0269] The sulfur-containing compound may be any suitable compound comprising one or more sulfur atoms. In one aspect, the sulfur-containing compound comprises a moiety selected from thionyl, thioether, sulfinyl, sulfonyl, sulfamide, sulfonamide, sulfonic acid, and sulfonic ester. Non-limiting examples of the sulfur-containing compounds include ethanethiol; 2-propanethiol; 2-propene-1-thiol; 2-mercaptoethanesulfonic acid; benzenethiol; benzene-1,2-dithiol; cysteine; methionine; glutathione; cystine; or a salt or solvate thereof.

[0270] In one aspect, an effective amount of such a compound described above to cellulosic material as a molar ratio to glucosyl units of cellulose is about 10^{-6} to about 10, e.g., about 10^{-6} to about 7.5, about 10^{-6} to about 5, about 10^{-6} to about 2.5, about 10^{-6} to about 1, about 10^{-5} to about 1, about 10^{-5} to about 10^{-1} , about 10^{-4} to about 10^{-1} , about 10^{-3} to about 10^{-1} , or about 10^{-3} to about 10^{-2} . In another aspect, an effective amount of such a compound described above is about 0.1 μ M to about 1 M, e.g., about 0.5 μ M to about 0.75 M, about 0.75 μ M to about 0.5 M, about 1 μ M to about 0.25 M, about 1 μ M to about 0.1 M, about 5 μ M to about 50 mM, about 10 μ M to about 25 mM, about 50 μ M to about 25 mM, about 10 μ M to about 10 mM, about 5 μ M to about 5 mM, or about 0.1 mM to about 1 mM.

[0271] The term "liquor" means the solution phase, either aqueous, organic, or a combination thereof, arising from treatment of a lignocellulose and/or hemicellulose material in a slurry, or monosaccharides thereof, e.g., xylose, arabinose, mannose, etc., under conditions as described herein, and the soluble contents thereof. A liquor for cellulolytic enhancement of a GH61 polypeptide can be produced by treating a lignocellulose or hemicellulose material (or feedstock) by applying heat and/or pressure, optionally in the presence of a catalyst, e.g., acid,

optionally in the presence of an organic solvent, and optionally in combination with physical disruption of the material, and then separating the solution from the residual solids. Such conditions determine the degree of cellulolytic enhancement obtainable through the combination of liquor and a GH61 polypeptide during hydrolysis of a cellulosic substrate by a cellulase preparation. The liquor can be separated from the treated material using a method standard in the art, such as filtration, sedimentation, or centrifugation.

[0272] In one aspect, an effective amount of the liquor to cellulose is about 10^{-6} to about 10 g per g of cellulose, e.g., about 10^{-6} to about 7.5 g, about 10^{-6} to about 5 g, about 10^{-6} to about 2.5 g, about 10^{-6} to about 1 g, about 10^{-5} to about 1 g, about 10^{-5} to about 10^{-1} g, about 10^{-4} to about 10^{-1} g, about 10^{-3} to about 10^{-1} g, or about 10^{-3} to about 10^{-2} g per g of cellulose.

[0273] In one aspect, the one or more (e.g., several) hemicellulolytic enzymes comprise a commercial hemicellulolytic enzyme preparation. Examples of commercial hemicellulolytic enzyme preparations suitable for use in the present invention include, for example, SHEARZYME™ (Novozymes A/S), CELLIC® HTec (Novozymes A/S), CELLIC® HTec2 (Novozymes A/S), CELLIC® HTec3 (Novozymes A/S), VISCOZYME® (Novozymes A/S), ULTRAFLO® (Novozymes A/S), PULPZYME® HC (Novozymes A/S), MULTIFECT® Xylanase (Genencor), ACCELLERASE® XY (Genencor), ACCELLERASE® XC (Genencor), ECOPULP® TX-200A (AB Enzymes), HSP 6000 Xylanase (DSM), DEPOL™ 333P (Biocatalysts Limit, Wales, UK), DEPOL™ 740L (Biocatalysts Limit, Wales, UK), and DEPOL™ 762P (Biocatalysts Limit, Wales, UK).

[0274] Examples of xylanases useful in the methods of the present invention include, but are not limited to, xylanases from *Aspergillus aculeatus* (GeneSeqP:AAR63790; WO 94/21785), *Aspergillus fumigatus* (WO 2006/078256), *Penicillium pinophilum* (WO 2011/041405), *Penicillium* sp. (WO 2010/126772), *Thielavia terrestris* NRRL 8126 (WO 2009/079210), and *Trichophaea saccata* GH10 (WO 2011/057083).

[0275] Examples of beta-xylosidases useful in the methods of the present invention include, but are not limited to, beta-xylosidases from *Neurospora crassa* (SwissProt accession number Q7SOW4), *Trichoderma reesei* (UniProtKB/TrEMBL accession number Q92458), and *Talaromyces emersonii* (SwissProt accession number Q8X212).

[0276] Examples of acetylxyylan esterases useful in the methods of the present invention include, but are not limited to, acetylxyylan esterases from *Aspergillus aculeatus* (WO 2010/108918), *Chaetomium globosum* (Uniprot accession number Q2GWX4), *Chaetomium gracile* (GeneSeqP accession number AAB82124), *Humicola insolens* DSM 1800 (WO 2009/073709), *Hypocrea jecorina* (WO 2005/001036), *Myceliophthera thermophila* (WO 2010/014880), *Neurospora crassa* (UniProt accession number q7s259), *Phaeosphaeria nodorum* (Uniprot accession number QOUHJ1), and *Thielavia terrestris* NRRL 8126 (WO 2009/042846).

[0277] Examples of feruloyl esterases (ferulic acid esterases) useful in the methods of the present invention include, but are not limited to, feruloyl esterases from *Humicola insolens* DSM 1800 (WO 2009/076122), *Neosartorya fischeri* (UniProt Accession number A1D9T4), *Neurospora crassa* (UniProt accession number Q9HGR3), *Penicillium aurantiogriseum* (WO 2009/127729), and *Thielavia terrestris* (WO 2010/053838 and WO 2010/065448).

[0278] Examples of arabinofuranosidases useful in the methods of the present invention include, but are not limited to, arabinofuranosidases from *Aspergillus niger* (GeneSeqP accession number AAR94170), *Humicola insolens* DSM 1800 (WO 2006/114094 and WO 2009/073383), and *M. giganteus* (WO 2006/114094).

[0279] Examples of alpha-glucuronidases useful in the methods of the present invention include, but are not limited to, alpha-glucuronidases from *Aspergillus clavatus* (UniProt accession number alcc12), *Aspergillus fumigatus* (SwissProt accession number Q4WW45), *Aspergillus niger* (UniProt accession number Q96WX9), *Aspergillus terreus* (SwissProt accession number Q0CJP9), *Humicola insolens* (WO 2010/014706), *Penicillium aurantiogriseum* (WO 2009/068565), *Talaromyces emersonii* (UniProt accession number Q8X211), and *Trichoderma reesei* (UniProt accession number Q99024).

[0280] The polypeptides having enzyme activity used in the methods of the present invention may be produced by fermentation of the above-noted microbial strains on a nutrient medium containing suitable carbon and nitrogen sources and inorganic salts, using procedures known in the art (see, e.g., Bennett, J.W. and LaSure, L. (eds.), *More Gene Manipulations in Fungi*, Academic Press, CA, 1991). Suitable media are available from commercial suppliers or may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection). Temperature ranges and other conditions suitable for growth and enzyme production are known in the art (see, e.g., Bailey, J.E., and Ollis, D.F., *Biochemical Engineering Fundamentals*, McGraw-Hill Book Company, NY, 1986).

[0281] The fermentation can be any method of cultivation of a cell resulting in the expression or isolation of an enzyme or protein. Fermentation may, therefore, be understood as comprising shake flask cultivation, or small- or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the enzyme to be expressed or isolated. The resulting enzymes produced by the methods described above may be recovered from the fermentation medium and purified by conventional procedures.

[0282] Fermentation. The fermentable sugars obtained from the hydrolyzed cellulosic material can be fermented by one or more (e.g., several) fermenting microorganisms capable of fermenting the sugars directly or indirectly into a desired fermentation product. "Fermentation" or "fermentation process" refers to any fermentation process or any process comprising a fermentation step. Fermentation processes also include fermentation processes used in the consumable alcohol industry (e.g., beer and wine), dairy industry (e.g., fermented dairy products), leather industry, and tobacco industry. The fermentation conditions depend on the

desired fermentation product and fermenting organism and can easily be determined by one skilled in the art.

[0283] In the fermentation step, sugars, released from the cellulosic material as a result of the pretreatment and enzymatic hydrolysis steps, are fermented to a product, e.g., ethanol, by a fermenting organism, such as yeast. Hydrolysis (saccharification) and fermentation can be separate or simultaneous, as described herein.

[0284] Any suitable hydrolyzed cellulosic material can be used in the fermentation step in practicing the present invention. The material is generally selected based on the desired fermentation product, i.e., the substance to be obtained from the fermentation, and the process employed, as is well known in the art.

[0285] The term "fermentation medium" is understood herein to refer to a medium before the fermenting microorganism(s) is(are) added, such as, a medium resulting from a saccharification process, as well as a medium used in a simultaneous saccharification and fermentation process (SSF).

[0286] "Fermenting microorganism" refers to any microorganism, including bacterial and fungal organisms, suitable for use in a desired fermentation process to produce a fermentation product. The fermenting organism can be hexose and/or pentose fermenting organisms, or a combination thereof. Both hexose and pentose fermenting organisms are well known in the art. Suitable fermenting microorganisms are able to ferment, i.e., convert, sugars, such as glucose, xylose, xylulose, arabinose, maltose, mannose, galactose, and/or oligosaccharides, directly or indirectly into the desired fermentation product.

[0287] Examples of bacterial and fungal fermenting organisms producing ethanol are described by Lin et al., 2006, Appl. Microbiol. Biotechnol. 69: 627-642.

[0288] Examples of fermenting microorganisms that can ferment hexose sugars include bacterial and fungal organisms, such as yeast. Preferred yeast includes strains of *Candida*, *Kluyveromyces*, and *Saccharomyces*, e.g., *Candida sonorensis*, *Kluyveromyces marxianus*, and *Saccharomyces cerevisiae*.

[0289] Examples of fermenting organisms that can ferment pentose sugars in their native state include bacterial and fungal organisms, such as some yeast. Preferred xylose fermenting yeast include strains of *Candida*, preferably *C. sheatae* or *C. sonorensis*; and strains of *Pichia*, preferably *P. stipitis*, such as *P. stipitis* CBS 5773. Preferred pentose fermenting yeast include strains of *Pachysolen*, preferably *P. tannophilus*. Organisms not capable of fermenting pentose sugars, such as xylose and arabinose, may be genetically modified to do so by methods known in the art.

[0290] Examples of bacteria that can efficiently ferment hexose and pentose to ethanol include, for example, *Bacillus coagulans*, *Clostridium acetobutylicum*, *Clostridium*

thermocellum, *Clostridium phytofermentans*, *Geobacillus* sp., *Thermoanaerobacter saccharolyticum*, and *Zymomonas mobilis* (Philippidis, 1996, *supra*).

[0291] Other fermenting organisms include strains of *Bacillus*, such as *Bacillus coagulans*; *Candida*, such as *C. sonorensis*, *C. methanosorbosa*, *C. diddensiae*, *C. parapsilosis*, *C. naedodendra*, *C. blankii*, *C. entomophila*, *C. brassicae*, *C. pseudotropicalis*, *C. boidinii*, *C. utilis*, and *C. scheidtiae*; *Clostridium*, such as *C. acetobutylicum*, *C. thermocellum*, and *C. phytofermentans*; *E. coli*, especially *E. coli* strains that have been genetically modified to improve the yield of ethanol; *Geobacillus* sp.; *Hansenula*, such as *Hansenula anomala*; *Klebsiella*, such as *K. oxytoca*; *Kluyveromyces*, such as *K. marxianus*, *K. lactis*, *K. thermotolerans*, and *K. fragilis*; *Schizosaccharomyces*, such as *S. pombe*; *Thermoanaerobacter*, such as *Thermoanaerobacter saccharolyticum*; and *Zymomonas*, such as *Zymomonas mobilis*.

[0292] In a preferred aspect, the yeast is a *Bretannomyces*. In a more preferred aspect, the yeast is *Bretannomyces clausenii*. In another preferred aspect, the yeast is a *Candida*. In another more preferred aspect, the yeast is *Candida sonorensis*. In another more preferred aspect, the yeast is *Candida boidinii*. In another more preferred aspect, the yeast is *Candida blankii*. In another more preferred aspect, the yeast is *Candida brassicae*. In another more preferred aspect, the yeast is *Candida diddensii*. In another more preferred aspect, the yeast is *Candida entomophiliia*. In another more preferred aspect, the yeast is *Candida pseudotropicalis*. In another more preferred aspect, the yeast is *Candida scheidtiae*. In another more preferred aspect, the yeast is *Candida utilis*. In another preferred aspect, the yeast is a *Clavispora*. In another more preferred aspect, the yeast is *Clavispora lusitaniae*. In another more preferred aspect, the yeast is *Clavispora opuntiae*. In another preferred aspect, the yeast is a *Kluyveromyces*. In another more preferred aspect, the yeast is *Kluyveromyces fragilis*. In another more preferred aspect, the yeast is *Kluyveromyces marxianus*. In another more preferred aspect, the yeast is *Kluyveromyces thermotolerans*. In another preferred aspect, the yeast is a *Pachysolen*. In another more preferred aspect, the yeast is *Pachysolen tannophilus*. In another preferred aspect, the yeast is a *Pichia*. In another more preferred aspect, the yeast is a *Pichia stipitis*. In another preferred aspect, the yeast is a *Saccharomyces* spp. In another more preferred aspect, the yeast is *Saccharomyces cerevisiae*. In another more preferred aspect, the yeast is *Saccharomyces distaticus*. In another more preferred aspect, the yeast is *Saccharomyces uvarum*.

[0293] In a preferred aspect, the bacterium is a *Bacillus*. In a more preferred aspect, the bacterium is *Bacillus coagulans*. In another preferred aspect, the bacterium is a *Clostridium*. In another more preferred aspect, the bacterium is *Clostridium acetobutylicum*. In another more preferred aspect, the bacterium is *Clostridium phytofermentans*. In another more preferred aspect, the bacterium is *Clostridium thermocellum*. In another more preferred aspect, the bacterium is *Geobacillus* sp. In another more preferred aspect, the bacterium is a *Thermoanaerobacter*. In another more preferred aspect, the bacterium is *Thermoanaerobacter saccharolyticum*. In another preferred aspect, the bacterium is a *Zymomonas*. In another more preferred aspect, the bacterium is *Zymomonas mobilis*.

[0294] Commercially available yeast suitable for ethanol production include, e.g., BIOFERM™ AFT and XR (NABC - North American Bioproducts Corporation, GA, USA), ETHANOL RED™ yeast (Fermentis/Lesaffre, USA), FALI™ (Fleischmann's Yeast, USA), FERMIOL™ (DSM Specialties), GERT STRAND™ (Gert Strand AB, Sweden), and SUPERSTART™ and THERMOSACC™ fresh yeast (Ethanol Technology, WI, USA).

[0295] In a preferred aspect, the fermenting microorganism has been genetically modified to provide the ability to ferment pentose sugars, such as xylose utilizing, arabinose utilizing, and xylose and arabinose co-utilizing microorganisms.

[0296] The cloning of heterologous genes into various fermenting microorganisms has led to the construction of organisms capable of converting hexoses and pentoses to ethanol (co-fermentation) (Chen and Ho, 1993, Cloning and improving the expression of *Pichia stipitis* xylose reductase gene in *Saccharomyces cerevisiae*, Appl. Biochem. Biotechnol. 39-40: 135-147; Ho et al., 1998, Genetically engineered *Saccharomyces* yeast capable of effectively cofermenting glucose and xylose, Appl. Environ. Microbiol. 64: 1852-1859; Kotter and Ciriacy, 1993, Xylose fermentation by *Saccharomyces cerevisiae*, Appl. Microbiol. Biotechnol. 38: 776-783; Walfridsson et al., 1995, Xylose-metabolizing *Saccharomyces cerevisiae* strains overexpressing the TKL1 and TAL1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase, Appl. Environ. Microbiol. 61: 4184-4190; Kuyper et al., 2004, Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle, FEMS Yeast Research 4: 655-664; Beall et al., 1991, Parametric studies of ethanol production from xylose and other sugars by recombinant *Escherichia coli*, Biotech. Bioeng. 38: 296-303; Ingram et al., 1998, Metabolic engineering of bacteria for ethanol production, Biotechnol. Bioeng. 58: 204-214; Zhang et al., 1995, Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*, Science 267: 240-243; Deanda et al., 1996, Development of an arabinose-fermenting *Zymomonas mobilis* strain by metabolic pathway engineering, Appl. Environ. Microbiol. 62: 4465-4470; WO 2003/062430, xylose isomerase).

[0297] In a preferred aspect, the genetically modified fermenting microorganism is *Candida sonorensis*. In another preferred aspect, the genetically modified fermenting microorganism is *Escherichia coli*. In another preferred aspect, the genetically modified fermenting microorganism is *Klebsiella oxytoca*. In another preferred aspect, the genetically modified fermenting microorganism is *Kluyveromyces marxianus*. In another preferred aspect, the genetically modified fermenting microorganism is *Saccharomyces cerevisiae*. In another preferred aspect, the genetically modified fermenting microorganism is *Zymomonas mobilis*.

[0298] It is well known in the art that the organisms described above can also be used to produce other substances, as described herein.

[0299] The fermenting microorganism is typically added to the degraded cellulosic material or hydrolysate and the fermentation is performed for about 8 to about 96 hours, e.g., about 24 to

about 60 hours. The temperature is typically between about 26°C to about 60°C, *e.g.*, about 32°C or 50°C, and about pH 3 to about pH 8, *e.g.*, pH 4-5, 6, or 7.

[0300] In one aspect, the yeast and/or another microorganism are applied to the degraded cellulosic material and the fermentation is performed for about 12 to about 96 hours, such as typically 24-60 hours. In another aspect, the temperature is preferably between about 20°C to about 60°C, *e.g.*, about 25°C to about 50°C, about 32°C to about 50°C, or about 32°C to about 50°C, and the pH is generally from about pH 3 to about pH 7, *e.g.*, about pH 4 to about pH 7. However, some fermenting organisms, *e.g.*, bacteria, have higher fermentation temperature optima. Yeast or another microorganism is preferably applied in amounts of approximately 10^5 to 10^{12} , preferably from approximately 10^7 to 10^{10} , especially approximately 2×10^8 viable cell count per ml of fermentation broth. Further guidance in respect of using yeast for fermentation can be found in, *e.g.*, "The Alcohol Textbook" (Editors K. Jacques, T.P. Lyons and D.R. Kelsall, Nottingham University Press, United Kingdom 1999).

[0301] A fermentation stimulator can be used in combination with any of the processes described herein to further improve the fermentation process, and in particular, the performance of the fermenting microorganism, such as, rate enhancement and ethanol yield. A "fermentation stimulator" refers to stimulators for growth of the fermenting microorganisms, in particular, yeast. Preferred fermentation stimulators for growth include vitamins and minerals. Examples of vitamins include multivitamins, biotin, pantothenate, nicotinic acid, meso-inositol, thiamine, pyridoxine, para-aminobenzoic acid, folic acid, riboflavin, and Vitamins A, B, C, D, and E. See, for example, Alfenore et al., Improving ethanol production and viability of *Saccharomyces cerevisiae* by a vitamin feeding strategy during fed-batch process, Springer-Verlag (2002). Examples of minerals include minerals and mineral salts that can supply nutrients comprising P, K, Mg, S, Ca, Fe, Zn, Mn, and Cu.

[0302] Fermentation products: A fermentation product can be any substance derived from the fermentation. The fermentation product can be, without limitation, an alcohol (*e.g.*, arabinitol, n-butanol, isobutanol, ethanol, glycerol, methanol, ethylene glycol, 1,3-propanediol [propylene glycol], butanediol, glycerin, sorbitol, and xylitol); an alkane (*e.g.*, pentane, hexane, heptane, octane, nonane, decane, undecane, and dodecane), a cycloalkane (*e.g.*, cyclopentane, cyclohexane, cycloheptane, and cyclooctane), an alkene (*e.g.* pentene, hexene, heptene, and octene); an amino acid (*e.g.*, aspartic acid, glutamic acid, glycine, lysine, serine, and threonine); a gas (*e.g.*, methane, hydrogen (H_2), carbon dioxide (CO_2), and carbon monoxide (CO)); isoprene; a ketone (*e.g.*, acetone); an organic acid (*e.g.*, acetic acid, acetonc acid, adipic acid, ascorbic acid, citric acid, 2,5-diketo-D-gluconic acid, formic acid, fumaric acid, glucaric acid, gluconic acid, glucuronic acid, glutaric acid, 3-hydroxypropionic acid, itaconic acid, lactic acid, malic acid, malonic acid, oxalic acid, oxaloacetic acid, propionic acid, succinic acid, and xylonic acid); and polyketide. The fermentation product can also be protein as a high value product.

[0303] In a preferred aspect, the fermentation product is an alcohol. It will be understood that

the term "alcohol" encompasses a substance that contains one or more hydroxyl moieties. In a more preferred aspect, the alcohol is n-butanol. In another more preferred aspect, the alcohol is isobutanol. In another more preferred aspect, the alcohol is ethanol. In another more preferred aspect, the alcohol is methanol. In another more preferred aspect, the alcohol is arabinitol. In another more preferred aspect, the alcohol is butanediol. In another more preferred aspect, the alcohol is ethylene glycol. In another more preferred aspect, the alcohol is glycerin. In another more preferred aspect, the alcohol is glycerol. In another more preferred aspect, the alcohol is 1,3-propanediol. In another more preferred aspect, the alcohol is sorbitol. In another more preferred aspect, the alcohol is xylitol. See, for example, Gong, C. S., Cao, N. J., Du, J., and Tsao, G. T., 1999, Ethanol production from renewable resources, in *Advances in Biochemical Engineering/Biotechnology*, Scheper, T., ed., Springer-Verlag Berlin Heidelberg, Germany, 65: 207-241; Silveira, M. M., and Jonas, R., 2002, The biotechnological production of sorbitol, *Appl. Microbiol. Biotechnol.* 59: 400-408; Nigam, P., and Singh, D., 1995, Processes for fermentative production of xylitol - a sugar substitute, *Process Biochemistry* 30 (2): 117-124; Ezeji, T. C., Qureshi, N. and Blaschek, H. P., 2003, Production of acetone, butanol and ethanol by *Clostridium beijerinckii* BA101 and in situ recovery by gas stripping, *World Journal of Microbiology and Biotechnology* 19 (6): 595-603.

[0304] In another preferred aspect, the fermentation product is an alkane. The alkane can be an unbranched or a branched alkane. In another more preferred aspect, the alkane is pentane. In another more preferred aspect, the alkane is hexane. In another more preferred aspect, the alkane is heptane. In another more preferred aspect, the alkane is octane. In another more preferred aspect, the alkane is nonane. In another more preferred aspect, the alkane is decane. In another more preferred aspect, the alkane is undecane. In another more preferred aspect, the alkane is dodecane.

[0305] In another preferred aspect, the fermentation product is a cycloalkane. In another more preferred aspect, the cycloalkane is cyclopentane. In another more preferred aspect, the cycloalkane is cyclohexane. In another more preferred aspect, the cycloalkane is cycloheptane. In another more preferred aspect, the cycloalkane is cyclooctane.

[0306] In another preferred aspect, the fermentation product is an alkene. The alkene can be an unbranched or a branched alkene. In another more preferred aspect, the alkene is pentene. In another more preferred aspect, the alkene is hexene. In another more preferred aspect, the alkene is heptene. In another more preferred aspect, the alkene is octene.

[0307] In another preferred aspect, the fermentation product is an amino acid. In another more preferred aspect, the organic acid is aspartic acid. In another more preferred aspect, the amino acid is glutamic acid. In another more preferred aspect, the amino acid is glycine. In another more preferred aspect, the amino acid is lysine. In another more preferred aspect, the amino acid is serine. In another more preferred aspect, the amino acid is threonine. See, for example, Richard, A., and Margaritis, A., 2004, Empirical modeling of batch fermentation kinetics for poly(glutamic acid) production and other microbial biopolymers, *Biotechnology and Bioengineering* 87 (4): 501-515.

[0308] In another preferred aspect, the fermentation product is a gas. In another more preferred aspect, the gas is methane. In another more preferred aspect, the gas is H₂. In another more preferred aspect, the gas is CO₂. In another more preferred aspect, the gas is CO. See, for example, Kataoka, N., A. Miya, and K. Kiriya, 1997, Studies on hydrogen production by continuous culture system of hydrogen-producing anaerobic bacteria, *Water Science and Technology* 36 (6-7): 41-47; and Gunaseelan V.N. in *Biomass and Bioenergy*, Vol. 13 (1-2), pp. 83-114, 1997, Anaerobic digestion of biomass for methane production: A review.

[0309] In another preferred aspect, the fermentation product is isoprene.

[0310] In another preferred aspect, the fermentation product is a ketone. It will be understood that the term "ketone" encompasses a substance that contains one or more ketone moieties. In another more preferred aspect, the ketone is acetone. See, for example, Qureshi and Blaschek, 2003, *supra*.

[0311] In another preferred aspect, the fermentation product is an organic acid. In another more preferred aspect, the organic acid is acetic acid. In another more preferred aspect, the organic acid is acetonetic acid. In another more preferred aspect, the organic acid is adipic acid. In another more preferred aspect, the organic acid is ascorbic acid. In another more preferred aspect, the organic acid is citric acid. In another more preferred aspect, the organic acid is 2,5-diketo-D-gluconic acid. In another more preferred aspect, the organic acid is formic acid. In another more preferred aspect, the organic acid is fumaric acid. In another more preferred aspect, the organic acid is glucaric acid. In another more preferred aspect, the organic acid is gluconic acid. In another more preferred aspect, the organic acid is glucuronic acid. In another more preferred aspect, the organic acid is glutaric acid. In another preferred aspect, the organic acid is 3-hydroxypropionic acid. In another more preferred aspect, the organic acid is itaconic acid. In another more preferred aspect, the organic acid is lactic acid. In another more preferred aspect, the organic acid is malic acid. In another more preferred aspect, the organic acid is malonic acid. In another more preferred aspect, the organic acid is oxalic acid. In another more preferred aspect, the organic acid is propionic acid. In another more preferred aspect, the organic acid is succinic acid. In another more preferred aspect, the organic acid is xylonic acid. See, for example, Chen, R., and Lee, Y. Y., 1997, Membrane-mediated extractive fermentation for lactic acid production from cellulosic biomass, *Appl. Biochem. Biotechnol.* 63-65: 435-448.

[0312] In another preferred aspect, the fermentation product is polyketide.

[0313] Recovery. The fermentation product(s) can be optionally recovered from the fermentation medium using any method known in the art including, but not limited to, chromatography, electrophoretic procedures, differential solubility, distillation, or extraction. For example, alcohol is separated from the fermented cellulosic material and purified by conventional methods of distillation. Ethanol with a purity of up to about 96 vol.% can be obtained, which can be used as, for example, fuel ethanol, drinking ethanol, *i.e.*, potable

neutral spirits, or industrial ethanol.

Paper-making

[0314] The present disclosure also relates to processes for manufacturing a paper material, which processes comprise treating a paper-making pulp and/or process water with a polypeptide having endoglucanase activity of the present invention.

[0315] The term "a paper-making process" means a process, wherein the pulp is suspended in water, mixed with various additives, and then passed to equipment in which paper, cardboard, tissue, towel etc. is formed, pressed and dried.

[0316] The term "paper material" means products, which can be made out of pulp, such as paper, linerboard, corrugated paperboard, tissue, towels, corrugated containers, or boxes.

[0317] The term "a paper-making pulp" or "pulp" means any pulp which can be used for the production of a paper material. For example, the pulp can be supplied as a virgin pulp, or can be derived from a recycled source. The paper-making pulp may be a wood pulp, a non-wood pulp or a pulp made from waste paper. A wood pulp may be made from softwood such as pine, redwood, fir, spruce, cedar, and hemlock or from hardwood such as maple, alder, birch, hickory, beech, aspen, acacia, and eucalyptus. A non-wood pulp may be made, e.g., from bagasse, bamboo, cotton, or kenaf. A waste paper pulp may be made by re-pulping waste paper such as newspaper, mixed office waste, computer print-out, white ledger, magazines, milk cartons, paper cups, etc.

[0318] In a particular aspect, the paper-making pulp to be treated comprises both hardwood pulp and softwood pulp.

[0319] The wood pulp to be treated may be mechanical pulp (such as ground wood pulp, GP), chemical pulp (such as Kraft pulp or sulfite pulp), semichemical pulp (SCP), thermomechanical pulp (TMP), chemithermomechanical pulp (CTMP), or bleached chemithermomechanical pulp (BCTMP).

[0320] Mechanical pulp is manufactured by grinding and refining methods, wherein the raw material is subjected to periodical pressure impulses. The following are examples of mechanical pulps: TMP is thermomechanical pulp, GW is groundwood pulp, PGW is pressurized groundwood pulp, RMP is refiner mechanical pulp, PRMP is pressurized refiner mechanical pulp, and CTMP is chemithermomechanical pulp.

[0321] Chemical pulp is manufactured by alkaline cooking whereby most of the lignin and hemicellulose components are removed. In Kraft pulping or sulphate cooking sodium sulphide or sodium hydroxide are used as principal cooking chemicals.

[0322] The Kraft pulp to be treated may be a bleached Kraft pulp, which may consist of softwood bleached Kraft (SWBK, also called NBKP, Nadel Holz Bleached Kraft Pulp), hardwood bleached Kraft (HWBK, also called LBKP, Laub Holz Bleached Kraft Pulp), or a mixture of these.

[0323] The pulp to be used in the processes of the present invention is a suspension of mechanical pulp, chemical pulp, or a combination thereof. For example, the pulp to be used in the processes of the present invention may comprise 0%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%, 60-70%, 70-80%, 80-90%, or 90-100% of chemical pulp. In a particular aspect, a chemical pulp forms part of the pulp being used for manufacturing the paper material. The expression "forms part of" means that the percentage of chemical pulp lies within the range of 1-99%. In particular aspects, the percentage of chemical pulp lies within the range of 2-98%, 3-97%, 4-96%, 5-95%, 6-94%, 7-93%, 8-92%, 9-91%, 10-90%, 15-85%, 20-80%, 25-75%, 30-70%, 40-60%, or 45-55%.

[0324] In a particular aspect, the chemical pulp is a Kraft pulp, a sulfite pulp, a semichemical pulp (SCP), a thermomechanical pulp (TMP), a chemithermomechanical pulp (CTMP), or a bleached chemithermomechanical pulp (BCTMP). In particular aspects, the Kraft pulp is bleached Kraft pulp, for example, softwood bleached Kraft (SWBK, also called NBKP, Nadel Holz Bleached Kraft Pulp), hardwood bleached Kraft (HWBK, also called LBKP, Laub Holz Bleached Kraft Pulp), or a mixture thereof.

[0325] In the case of pulp and paper processing, the processes of the present invention can be carried out at any pulp production stage. The enzyme can be added to any holding tank, e.g., to a pulp storing container (storage chest), storage tower, mixing chest, or metering chest. The enzyme treatment can be performed before the bleaching of pulp, in connection with the pulp bleaching process, or after the bleaching. When carried out in connection with pulp bleaching the enzyme preparation may be added together with bleaching chemicals such as chlorine or chlorine dioxide.

[0326] Oxygen gas, hydrogen peroxide, ozone, or combinations thereof may also be applied in the bleaching of pulp. The enzyme preparation may also be added together with these substances.

[0327] Preferably the enzyme preparation is added prior to bleaching. The enzyme can also be added to the circulated process water (white water) originating from bleaching and process water (brown water) originating from the mechanical or chemimechanical pulping process. In a particular aspect of a Kraft pulping process, the enzyme is added during the brown-stock washing.

[0328] The term "process water" comprises (1) water added as a raw material to the paper manufacturing process; (2) intermediate water products resulting from any step of the process for manufacturing the paper material; and/or (3) waste water as an output or by-product of the process. In a particular aspect, the process water is, has been, is being, or is intended for

being circulated (re-circulated), *i.e.*, re-used in another step of the process. The term "water" means any aqueous medium, solution, suspension, *e.g.*, ordinary tap water, and tap water in admixture with various additives and adjuvants commonly used in paper manufacturing processes. In a particular aspect, the process water has a low content of solid (dry) matter, *e.g.*, below 20%, 18%, 16%, 14%, 12%, 10%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1% dry matter.

[0329] The processes of the present invention may be carried out under conventional conditions in the pulp and paper processing. The process conditions will be a function of the enzyme(s) applied, the reaction time, and the conditions used.

[0330] The drainability of paper-making pulps may also be improved by treatment of the pulp with a polypeptide of the present invention. Use of a polypeptide of the present invention may be more effective, *e.g.*, result in a higher degree of loosening bundles of strongly hydrated microfibrils in the fines fraction that limits the rate of drainage by blocking hollow spaces between the fibers and in the wire mesh of the paper machine.

[0331] A polypeptide having endoglucanase activity of the present invention should be added in an effective amount. The term "effective amount" means the amount sufficient to achieve the desired and expected effect, such as (a) an increase in handsheet tensile index, and/or (b) an increase in handsheet tear index, relative to no treatment with the polypeptide having endoglucanase activity.

[0332] In a particular aspect, the dosage of a polypeptide of the present invention and additional enzymes, if any, is from about 0.1 mg enzyme protein to about 100,000 mg enzyme protein (of each enzyme) per ton of paper pulp.

[0333] In further particular aspects, the amount of the polypeptide having endoglucanase activity and additional enzymes, if any, is in the range of 0.00001-20 mg of enzyme (calculated as pure enzyme protein) per gram (dry weight) of lignocellulosic material, *e.g.*, 0.0001-20, 0.0001-10, 0.0001-1, 0.001-1, 0.001-0.1, or 0.01-0.1 mg of enzyme per gram of lignocellulosic material. These amounts refer to the amount of each enzyme.

[0334] The enzymatic treatment can be performed at conventional consistency, *e.g.*, 0.5-10 % dry substance. In particular aspects, the consistency is within the range of 0.5-45, 0.5-40, 0.5-35, 0.5-30, 0.5-25, 0.5-20, 0.5-15, 0.5-10, 0.5-8, 0.5-6, or 0.5-5% dry substance.

[0335] The enzymatic treatment may be carried out at a temperature of from about 10 to about 100°C. Further examples of temperature ranges (all "from about" and "to about") are the following: 20-100, 30-100, 35-100, 37-100, 40-100, 50-100, 60-100, 70-100, 10-90, 10-80, 10-70, 10-60, or 30-60°C, as well as any combination of the upper and lower values indicated herein. A typical temperature is from about 20 to 90°C, or 20 to 95°C, preferably from about 40 to 70°C, or 40 to 75°C.

[0336] The enzymatic treatment may be carried out at a pH from about 2 to about 12. Further

examples of pH ranges (all "from about" and "to about") are the following: 3-12, 4-12, 5-12, 6-12, 7-12, 8-12, 9-12, 2-11, 2-10, 2-9, 2-8, 4-10, or 5-8 as well as any combination of the upper and lower values indicated herein. A typical pH range is from about 2 to 11, preferably within the range from about 4 to 9.5, or 6 to 9.

[0337] A suitable duration of the enzymatic treatment may be in the range from a few seconds to several hours, *e.g.*, about 30 seconds to about 48 hours, about 1 minute to about 24 hours, about 1 minute to about 18 hours, about 1 minute to about 12 hours, about 1 minute to 5 hours, about 1 minute to about 2 hours, about 1 minute to about 1 hour, or about 1 minute to about 30 minutes. A typical reaction time is from about 10 minutes to about 3 hours, about 10 minutes to about 10 hours, preferably about 15 minutes to about 1 hour, or about 15 minutes to about 2 hours.

[0338] Molecular oxygen from the atmosphere will usually be present in sufficient quantity, if required. Therefore, the reaction may be conveniently performed in an open reactor, *i.e.*, at atmospheric pressure.

[0339] Various additives over and above a polypeptide having endoglucanase activity of the present invention and additional enzymes, if any, can be used in the processes of the present invention. Surfactants and/or dispersants are often present in, and/or added to, a paper-making pulp. Thus the processes of the present invention may be carried out in the presence of an anionic, non-ionic, cationic, and/or zwitterionic surfactant and/or dispersant conventionally used in a paper-making pulp. Examples of anionic surfactants are carboxylates, sulphates, sulphonates, or phosphates of alkyl, substituted alkyl, or aryl. Fatty acids are examples of alkyl-carboxylates. Examples of non-ionic surfactants are polyoxyethylene compounds, such as alcohol ethoxylates, propoxylates, or mixed ethoxy/propoxylates, polyglycerols and other polyols, as well as certain block-copolymers. Examples of cationic surfactants are water-soluble cationic polymers, such as quaternary ammonium sulphates and certain amines, *e.g.*, epichlorohydrin/dimethylamine polymers (EPI-DMA) and cross-linked solutions thereof, polydiallyl dimethyl ammonium chloride (DADMAC), DADMAC/acrylamide copolymers, and ionene polymers, such as those disclosed in U.S. Patent Nos. 5,681,862 and 5,575,993. Examples of zwitterionic or amphoteric surfactants are betaines, glycinate, amino propionates, imino propionates, and various imidazoline-derivatives. Also the polymers disclosed in U.S. Patent No. 5,256,252 may be used.

[0340] Also according to the present disclosure, surfactants such as the above, including any combination thereof, may be used in a paper making process together with a polypeptide having endoglucanase activity of the present invention, and included in a composition together with such enzyme. The amount of each surfactant in such composition may amount to about 8 to about 40% (w/w) of the composition. In particular aspects, the amount of each surfactant is from about 10 to about 38, about 12 to about 36, about 14 to about 34, about 16 to about 34, about 18 to about 34, about 20 to about 34, about 22 to about 34, about 24 to about 34, about 26 to about 34, or about 28 to about 32% (w/w).

[0341] In another particular embodiment, each of the above ranges refers to the total amount of surfactants.

[0342] In the processes of the present disclosure, a polypeptide of the present invention may be applied alone or together with an additional enzyme. The term "an additional enzyme" means at least one additional enzyme, e.g., one, two, three, four, five, six, seven, eight, nine, ten or even more additional enzymes.

[0343] The term "applied together with" (or "used together with") means that the additional enzyme may be applied in the same step or in another step of the process. The other process step may be upstream or downstream in the paper manufacturing process, as compared to the step in which the paper making pulp or process water is treated with a polypeptide of the present invention.

[0344] In particular aspects, the additional enzyme is an enzyme which has protease, lipase, xylanase, cutinase, oxidoreductase, cellulase, endoglucanase, amylase, mannanase, steryl esterase, fatty acid oxidizing enzyme, and/or cholesterol esterase activity. Examples of oxidoreductase enzymes are enzymes with laccase and/or peroxidase activity. In a preferred aspect, the additional enzyme is lipase.

[0345] The term "a step" of a process means at least one step, and it could be one, two, three, four, five or even more process steps. In other words, a polypeptide of the present invention may be applied in at least one process step, and the additional enzyme(s) may also be applied in at least one process step, which may be the same or a different process step compared to the step where the polypeptide of the present invention is used.

[0346] The term "enzyme preparation" means a product containing at least one polypeptide of the present invention. The enzyme preparation may also comprise enzymes having other enzyme activities, preferably lipolytic enzymes or enzymes having oxidoreductase activity, most preferably lipolytic enzymes. In addition to the enzymatic activity such a preparation preferably contains at least one adjuvant. Examples of adjuvants used in enzyme preparations for the pulp and paper industry are buffers, polymers, surfactants, and stabilizing agents.

[0347] Any enzyme having protease, lipase, xylanase, cutinase, oxidoreductase, cellulase, endoglucanase, amylase, mannanase, steryl esterase, fatty acid oxidizing enzyme, and/or cholesterol esterase can be used as additional enzymes in the processes of the present invention. Non-limiting examples are listed below of such additional enzymes. The activity of any of the additional enzymes can be analyzed using any method known in the art for the enzyme in question, including the methods mentioned in the references cited.

[0348] Examples of cutinases are those derived from *Humicola insolens* (U.S. Patent No. 5,827,719); or a strain of *Fusarium*, e.g., *F. roseum culmorum*, or particularly *F. solani pisi* (WO 90/09446; WO 94/14964, WO 94/03578). The cutinase may also be derived from a strain of *Rhizoctonia*, e.g., *R. solani*, or a strain of *Alternaria*, e.g., *A. brassicicola* (WO 94/03578), or

variants thereof such as those described in WO 00/34450, or WO 01/92502.

[0349] Examples of proteases are the ALCALASE™, ESPERASE™, SAVINASE™, NEUTRASE™, and DURAZYM™ proteases (Novozymes A/S, Bagsvaerd, Denmark). Other proteases are derived from *Nocardiosis*, *Aspergillus*, *Rhizopus*, *Bacillus alcalophilus*, *B. cereus*, *B. naffo*, *B. vulgatus*, *B. mycoide*, especially proteases from the species *Nocardiosis* sp. and *Nocardiosis dassonvillei* such as those disclosed in WO 88/03947, and mutants thereof, e.g., those disclosed in WO 91/00345 and EP 415296.

[0350] Examples of amylases are the BAN™, AQUAZYM™, TERMAMYL™, and AQUAZYM™ Ultra amylases (Novozymes A/S, Bagsvaerd, Denmark). An example of a lipase is RESINASE™ A2X lipase (Novozymes A/S, Bagsvaerd, Denmark). An example of a xylanase is the PULPZYME™ HC hemicellulase (Novozymes A/S, Bagsvaerd, Denmark). Examples of other endoglucanases are the NOVOZYM 613, 342, and 476 enzyme products (Novozymes A/S, Bagsvaerd, Denmark).

[0351] Examples of mannanases are the *Trichoderma reesei* endo-beta-mannanases described by Stahlband et al., 1993, J. Biotechnol. 29: 229-242.

[0352] Examples of steryl esterases, peroxidases, laccases, and cholesterol esterases are disclosed in WO 00/53843; U.S. Patent No. 6,066,486; JP 2000080581; Karlsson et al., 2001, Appl. Microbiol. Biotechnol. 55: 317-320; and Zhang, 2000; Pulp & Paper Canada 101: 59-62. Further examples of oxidoreductases are the peroxidases and laccases disclosed in EP 730641; WO 01/98469; EP 719337; EP 765394; EP 767836; EP 763115; and EP 788547. Whenever an oxidoreductase enzyme is mentioned that requires or benefits from the presence of acceptors (e.g., oxygen or hydrogen peroxide), enhancers, mediators and/or activators, such compounds should be considered to be included. Examples of enhancers and mediators are disclosed in EP 705327; WO 98/56899; EP 677102; EP 781328; and EP 707637. If desired a distinction could be made by defining an oxidoreductase enzyme system (e.g., a laccase, or a peroxidase enzyme system) as the combination of the enzyme in question and its acceptor, and optionally also an enhancer and/or mediator for the enzyme in question.

[0353] Examples of fatty acid oxidizing enzymes are disclosed in WO 2003/035972.

[0354] The present invention is further described by the following examples.

Examples

Materials

[0355] Chemicals used as buffers and substrates were commercial products of at least

reagent grade.

Strains

[0356] *Chaetomium virescens* ATCC 32319 was used as the source of a polypeptide having endoglucanase activity. *Aspergillus oryzae* MT3568 strain was used for expression of the *Chaetomium virescens* gene encoding the polypeptide having endoglucanase activity. *A. oryzae* MT3568 is an *amdS* (acetamidase) disrupted gene derivative of *Aspergillus oryzae* JaL355 (WO 2002/40694) in which *pyrG* auxotrophy was restored by disrupting the *A. oryzae* acetamidase (*amdS*) gene.

Media and Solutions

[0357] YP+2% glucose medium was composed of 1% yeast extract, 2% peptone, and 2% glucose.

[0358] PDA plates were composed of potato infusion made by boiling 300 g of sliced (washed but unpeeled) potatoes in water for 30 minutes and then decanting or straining the broth through cheesecloth. Distilled water was then added until the total volume of the suspension was one liter, followed by 20 g of dextrose and 20 g of agar powder. The medium was sterilized by autoclaving at 15 psi for 15 minutes (Bacteriological Analytical Manual, 8th Edition, Revision A, 1998).

[0359] LB medium was composed of 10 g of Bacto-Tryptone, 5 g of yeast extract, 10 g of sodium chloride, and deionized water to 1 liter.

[0360] LB plates were composed of 10 g of Bacto-Tryptone, 5 g of yeast extract, 10 g of sodium chloride, 15 g of Bacto-agar, and deionized water to 1 liter. The medium was sterilized by autoclaving at 15 psi for 15 minutes (Bacteriological Analytical Manual, 8th Edition, Revision A, 1998).

[0361] COVE sucrose plates were composed of 342 g of sucrose, 20 g of agar powder, 20 ml of COVE salt solution, and deionized water to 1 liter. The medium was sterilized by autoclaving at 15 psi for 15 minutes (Bacteriological Analytical Manual, 8th Edition, Revision A, 1998). The medium was cooled to 60°C and 10 mM acetamide, 15 mM CsCl, and TRITON® X-100 (50 µl/500 ml) were added.

[0362] COVE salt solution was composed of 26 g of MgSO₄·7H₂O, 26 g of KCl, 26 g of KH₂PO₄, 50 ml of COVE trace metals solution, and deionized water to 1 liter.

[0363] COVE trace metals solution was composed of 0.04 g of Na₂B₄O₇·10H₂O, 0.4 g of

$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 1.2 g of $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0.7 g of $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0.8 g of $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 10 g of $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, and deionized water to 1 liter.

[0364] Dap-4C medium was composed of 20 g of dextrose, 10 g of maltose, 11 g of $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1 g of KH_2PO_4 , 2 g of citric acid, 5.2 g of $\text{K}_3\text{PO}_4 \cdot \text{H}_2\text{O}$, 0.5 g of yeast extract (Difco), 1 ml of Dowfax 63N10 (Dow Chemical Company), 0.5 ml of KU6 trace metals solution, 2.5 g of CaCO_3 , and deionized water to 1 liter. The medium was sterilized by autoclaving at 15 psi for 15 minutes (Bacteriological Analytical Manual, 8th Edition, Revision A, 1998). Before use, 3.5 ml of sterile 50% $(\text{NH}_4)_2\text{HPO}_4$ and 5 ml of sterile 20% lactic acid were added per 150 ml of Dap-4C medium.

[0365] KU6 trace metals solution was composed of 0.13 g of NiCl_2 , 2.5 g of $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 13.9 g of $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 8.45 g of $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 6.8 g of ZnCl_2 , 3 g of citric acid, and deionized water to 1 liter.

[0366] MDU2BP medium was composed of 45 g of maltose, 1 g of $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1 g of NaCl , 2 g of K_2SO_4 , 12 g of KH_2PO_4 , 7 g of yeast extract, 2 g of urea, 0.5 ml of AMG trace metals solution, and deionized water to 1 liter; pH 5.0.

[0367] AMG trace metals solution was composed of 14.3 g of $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 2.5 g of $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0.5 g of $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$, 13.8 g of $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 8.5 g of $\text{MnSO}_4 \cdot 7\text{H}_2\text{O}$, 3 g of citric acid, and deionized water to 1 liter.

[0368] 2XYT plates were composed of 16 g of tryptone, 10 g of yeast extract, 5 g of NaCl , 15 g of Noble agar, and deionized water to 1 liter.

Example 1: Preparation of *Chaetomium virescens* ATCC 32319 Cel5 endoglucanase

[0369] The *Chaetomium virescens* ATCC 32319 GH5 endoglucanase gene (SEQ ID NO: 1 [DNA sequence] and SEQ ID NO: 2 [deduced amino acid sequence]) was cloned into an *Aspergillus oryzae* expression vector as described below.

[0370] Two synthetic oligonucleotide primers, shown below, were designed to PCR amplify the endoglucanase gene from *C. virescens* ATCC 32319 genomic DNA. Genomic DNA was isolated using a FASTDNA® Spin for Soil Kit (MP Biomedicals, OH, USA).

Primer F-Cv1:

5'-ACACAACTGGGGATCCACCATGCCTTCAGTCGTCCTTTCTC-3' (SEQ ID NO: 3)

Primer R-Cv1:

5'-CCCTCTAGATCTCGAGCAACAATGCCGACACACTCCA-3' (SEQ ID NO: 4)

[0371] Bold letters represent gene sequence. The underlined sequence is homologous to the insertion sites of plasmid pDau109 (WO 2005/042735).

[0372] The amplification reaction was composed of 1 µl of *C. virescens* ATCC 32319 genomic DNA, 5 µl of 5X PHUSION™ HF Buffer (Finnzymes, Espoo, Finland), 0.5 µl (2 Units/µl) of PHUSION™ High-Fidelity DNA Polymerase (Finnzymes, Espoo, Finland), 1 µl of 5 µM primer F-Cv1, 1 µl of 5 µM primer R-Cv1, 0.5 µl of 10 mM dNTP, and 16 µl of H₂O. The amplification reaction was incubated in a PTC-200 DNA ENGINE™ Thermal Cycler (MJ Research Inc., Waltham, MA, USA) programmed for 1 cycle at 95°C for 2 minutes; and 35 cycles each at 98°C for 10 seconds, 60°C for 30 seconds, and 72°C for 2.5 minutes.

[0373] A 1.34 kbp PCR reaction product was isolated by 1% agarose gel electrophoresis using 40 mM Tris base-20 mM sodium acetate-1 mM disodium EDTA (TAE) buffer and staining with SYBR® Safe DNA Gel Stain (Invitrogen Corp., Carlsbad, CA, USA). The DNA band was visualized with the aid of an EAGLE EYE® Imaging System (Stratagene, La Jolla, CA, USA) and a DARKREADER® Transilluminator (Clare Chemical Research, Dolores, CO, USA). The 1.34 kbp DNA band was excised from the gel and purified using a GFX® PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences, Piscataway, NJ, USA) according to the manufacturer's instructions. The 1.34 kbp fragment was cleaved with *Bam* HI and *Xho* I and purified using a GFX® PCR DNA and Gel Band Purification Kit according to the manufacturer's instructions. An IN-FUSION™ Cloning Kit (BD Biosciences, Palo Alto, CA, USA) was used to clone the fragment directly into the expression vector pDau109 digested with *Bam* HI and *Xho* I.

[0374] The cloning mixture was transformed into *E. coli* T0P10F competent cells (Invitrogen Corp., Carlsbad, CA, USA) according to the manufacturer's instructions. The transformation mixture was plated onto LB plates supplemented with 100 µg of ampicillin per ml. Plasmid minipreps were prepared from several transformants and sequenced. One plasmid with the correct *C. virescens* GH5 coding sequence (SEQ ID NO: 1) was chosen. The plasmid was designated pDAU222-PE03860005710 (Figure 1). Cloning of the *Chaetomium virescens* GH5 gene into *Bam* HI-*Xho* I digested pDau109 resulted in transcription of the *C. virescens* GH5 gene under the control of a NA2-tpi double promoter. NA2-tpi is a modified promoter from the *Aspergillus niger* neutral alpha-amylase gene in which the untranslated leader has been replaced by the untranslated leader from the *Aspergillus nidulans* triose phosphate isomerase gene.

[0375] The expression plasmid pDAU222-PE03860005710 was transformed into protoplasts of *Aspergillus oryzae* MT3568 according to the method of European Patent EP 0238023, pages 14-15. *Aspergillus oryzae* MT3568 is an *amdS* (acetamidase) disrupted derivative of *A. oryzae* JaL355 (WO 2002/40694) in which *pyrG* auxotrophy was restored in the process of inactivating the *A. oryzae amdS* gene.

[0376] Transformants were purified on COVE sucrose plates through single conidia prior to sporulating them on PDA plates. Production of the *Chaetomium virescens* GH5 polypeptide by the transformants was analyzed from culture supernatants of 1 ml 96 deep well stationary cultivations at 26°C in YP+2% glucose medium. Expression was verified by SDS-PAGE using a NU-PAGE® 10% Bis-Tris SDS-PAGE gel (Invitrogen, Carlsbad, CA, USA) and Coomassie blue staining where a band of approximately 50 kDa was observed. One transformant was selected and designated *Aspergillus oryzae* 17.4.

[0377] For larger scale production, *A. oryzae* 17.4 spores were spread onto a PDA plate and incubated for five days at 37°C. The confluent spore plate was washed twice with 5 ml of 0.01% TWEEN® 20 to maximize the number of spores collected. The spore suspension was then used to inoculate twenty-five 500 ml flasks containing 100 ml of Dap-4C medium. The cultures were incubated at 30°C with constant shaking at 85 rpm. At day four post-inoculation, the culture broths were collected by filtration through a bottle top MF75™ SUPOR® MachV 0.2 µm PES filter (Thermo Fisher Scientific, Roskilde, Denmark). The culture broths from this transformant produced a band of GH5 protein of approximately 50 kDa (designated EXP03745). The identity of this band as the *C. virescens* GH5 polypeptide was verified by peptide sequencing.

Alternative method for producing Cv1 GH5 from *Chaetomium virescens*.

[0378] Based on the nucleotide sequence identified as SEQ ID NO: 1, a synthetic gene can be obtained from a number of vendors such as Gene Art (GENEART AG BioPark, Josef-Engert-Str. 11, 93053, Regensburg, Germany) or DNA2.0 (DNA2.0, 1430 O'Brien Drive, Suite E, Menlo Park, CA 94025, USA). The synthetic gene can be designed to incorporate additional DNA sequences such as restriction sites or homologous recombination regions to facilitate cloning into an expression vector.

[0379] Using the two synthetic oligonucleotide primers F-Cv1 and R-Cv1 described above, a simple PCR reaction can be used to amplify the full-length open reading frame from the synthetic gene of SEQ ID NO: 1. The gene can then be cloned into an expression vector, for example as described above, and expressed in a host cell, for example in *Aspergillus oryzae* as described above.

Example 2: Characterization of the *Chaetomium virescens* GH5 endoglucanase gene

[0380] The genomic DNA sequence and deduced amino acid sequence of the *Chaetomium virescens* GH5 endoglucanase encoding sequence are shown in SEQ ID NO: 1 and SEQ ID NO: 2, respectively. The genomic DNA sequence of 1320 bp (including the stop codon) contains one intron located at nucleotides 1119 to 1178 of SEQ ID NO: 1. The genomic DNA sequence encodes a polypeptide of 419 amino acids. The %G+C content of the mature

polypeptide coding sequence is 58.7%. Using the SignalP software program (Nielsen et al., 1997, Protein Engineering 10: 1-6), a signal peptide of 17 residues was predicted. The predicted mature protein contains 402 amino acids with a predicted molecular mass of 43.1 kDa and a predicted isoelectric point of 5.5. The protein contains a cellulose binding module of the CBM1 type at the N-terminus (amino acids 23 to 58 of SEQ ID NO: 2). The catalytic domain is amino acids 93 to 419.

[0381] A comparative alignment of mature endoglucanase amino acid sequences, without the signal peptides, was determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-453) as implemented in the Needle program of EMBOSS with gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 matrix. The alignment showed that the deduced amino acid sequence of the *Chaetomium virescens* GH5 endoglucanase (mature polypeptide) shares 75.4% sequence identity (excluding gaps) to the deduced amino acid sequence of an endoglucanase from *Hypocrea jecorina* (GENESEQP:ASQ45591).

Example 3: Purification of *Chaetomium virescens* GH5 polypeptide having endoglucanase activity

[0382] Filtered broth (Example 1) was adjusted to pH 7.0 and filtered using a 0.22 µm PES filter (Thermo Fisher Scientific, Roskilde, Denmark). To the filtrate was added ammonium sulphate to 1.0 M pH 7.0. The filtrate was loaded onto a PHENYL SEPHAROSE™ 6 Fast Flow column (high sub) (GE Healthcare, Piscataway, NJ, USA) equilibrated with 1.0 M ammonium sulphate pH 7.0, and bound proteins were eluted with 25 mM HEPES pH 7.0. The fractions were pooled and applied to a SEPHADEX™ G-25 (medium) column (GE Healthcare, Piscataway, NJ, USA) equilibrated in 50 mM HEPES pH 7.0. The fractions were applied to a SOURCE™ 15Q column (GE Healthcare, Piscataway, NJ, USA) equilibrated in 50 mM HEPES pH 7.0. The protein eluted in the flow-through. The flow-through was concentrated 5-fold using a VIVASPIN® 20 centrifugal concentrator with a 10,000 MWCO PES membrane (Sartorius Stedim Biotech GmbH, 37070 Goettingen, Germany).

Example 4: Pretreated corn stover hydrolysis assay

[0383] Corn stover was pretreated at the U.S. Department of Energy National Renewable Energy Laboratory (NREL) using 1.4 wt % sulfuric acid at 165°C and 107 psi for 8 minutes. The water-insoluble solids in the pretreated corn stover (PCS) contained 56.5% cellulose, 4.6% hemicellulose, and 28.4% lignin. Cellulose and hemicellulose were determined by a two-stage sulfuric acid hydrolysis with subsequent analysis of sugars by high performance liquid chromatography using NREL Standard Analytical Procedure #002. Lignin was determined gravimetrically after hydrolyzing the cellulose and hemicellulose fractions with sulfuric acid using NREL Standard Analytical Procedure #003.

[0384] Unmilled, unwashed PCS (whole slurry PCS) was prepared by adjusting the pH of PCS to 5.0 by addition of 10 M NaOH with extensive mixing, and then autoclaving for 20 minutes at 120°C. The dry weight of the whole slurry PCS was 29%. Milled unwashed PCS (dry weight 32.35%) was prepared by milling whole slurry PCS in a Cosmos ICMG 40 wet multi-utility grinder (EssEmm Corporation, Tamil Nadu, India).

[0385] The hydrolysis of PCS was conducted using 2.2 ml deep-well plates (Axygen, Union City, CA, USA) in a total reaction volume of 1.0 ml. The hydrolysis was performed with 50 mg of insoluble PCS solids per ml of 50 mM sodium acetate pH 5.0 buffer containing 1 mM manganese sulfate and various protein loadings of various enzyme compositions (expressed as mg protein per gram of cellulose). Enzyme compositions were prepared and then added simultaneously to all wells in a volume ranging from 50 µl to 200 µl, for a final volume of 1 ml in each reaction. The plate was then sealed using an ALPS-300™ plate heat sealer (Abgene, Epsom, United Kingdom), mixed thoroughly, and incubated at a specific temperature for 72 hours. All experiments reported were performed in triplicate.

[0386] Following hydrolysis, samples were filtered using a 0.45 µm MULTISCREEN® 96-well filter plate (Millipore, Bedford, MA, USA) and filtrates analyzed for sugar content as described below. When not used immediately, filtered aliquots were frozen at -20°C. The sugar concentrations of samples diluted in 0.005 M H₂SO₄ were measured using a 4.6 x 250 mm AMINEX® HPX-87H column (Bio-Rad Laboratories, Inc., Hercules, CA, USA) by elution with 0.05% w/w benzoic acid-0.005 M H₂SO₄ at 65°C at a flow rate of 0.6 ml per minute, and quantitation by integration of the glucose, cellobiose, and xylose signals using refractive index detection (CHEMSTATION®, AGILENT® 1100 HPLC, Agilent Technologies, Santa Clara, CA, USA) calibrated by pure sugar samples. The resultant glucose and cellobiose equivalents were used to calculate the percentage of cellulose conversion for each reaction.

[0387] Glucose, cellobiose, and xylose were measured individually. Measured sugar concentrations were adjusted for the appropriate dilution factor. In case of unwashed PCS, the net concentrations of enzymatically-produced sugars were determined by adjusting the measured sugar concentrations for corresponding background sugar concentrations in unwashed PCS at zero time. All HPLC data processing was performed using MICROSOFT EXCEL™ software (Microsoft, Richland, WA, USA).

[0388] The degree of cellulose conversion to glucose was calculated using the following equation: % conversion = (glucose concentration / glucose concentration in a limit digest) x 100. To calculate total conversion the glucose and cellobiose values were combined. Cellobiose concentration was multiplied by 1.053 in order to convert to glucose equivalents and added to the glucose concentration. The degree of total cellulose conversion was calculated using the following equation: % conversion = [(glucose concentration + 1.053 x (cellobiose concentration)) / ((glucose concentration + 1.053 x (cellobiose concentration) in a limit digest)] x 100. The 1.053 factor for cellobiose takes into account the increase in mass when cellobiose is converted to glucose. In order to calculate % conversion, a 100% conversion point was set

based on a cellulase control (100 mg of *Trichoderma reesei* cellulase per gram cellulose), and all values were divided by this number and then multiplied by 100. Triplicate data points were averaged and standard deviation was calculated.

Example 5: Preparation of *Aspergillus fumigatus* NN055679 Cel7A cellobiohydrolase I

[0389] A tfasty search (Pearson et al., 1997, Genomics 46:24-36) of the *Aspergillus fumigatus* partial genome sequence (The Institute for Genomic Research, Rockville, MD) was performed using as query a Cel7 cellobiohydrolase protein sequence from *Trichoderma reesei* (Accession No. P00725). Several genes were identified as putative Family GH7 homologs based upon a high degree of similarity to the query sequence at the amino acid level. One genomic region with significant identity to the query sequence was chosen, and the corresponding gene was designated *cel7A*.

[0390] Two synthetic oligonucleotide primers shown below were designed to PCR amplify the *Aspergillus fumigatus* NN055679 *cel7A* cellobiohydrolase I gene (SEQ ID NO: 5 [DNA sequence] and SEQ ID NO: 6 [deduced amino acid sequence]) from genomic DNA of *Aspergillus fumigatus* prepared as described in WO 2005/047499.

Forward primer:

5'-gggcATGCTGGCCTCCACCTTCTCC-3' (SEQ ID NO: 7)

Reverse primer:

5'-gggtaattaaCTACAGGCACTGAGAGTAA-3' (SEQ ID NO: 8)

[0391] Upper case letters represent the coding sequence. The remainder of the sequence provides restriction endonuclease sites for *Sph* I and *Pac* I in the forward and reverse sequences, respectively. Using these primers, the *Aspergillus fumigatus cel7A* gene was amplified using standard PCR methods and the reaction product isolated by 1% agarose gel electrophoresis using TAE buffer and purified using a QIAQUICK® Gel Extraction Kit (QIAGEN Inc., Valencia, CA, USA) according to the manufacturer's instructions.

[0392] The fragment was digested with *Sph* I and *Pac* I and ligated into the expression vector pAIIo2 (WO 2004/099228) also digested with *Sph* I and *Pac* I according to standard procedures. The ligation products were transformed into *E. coli* XL10 SOLOPACK® cells (Stratagene, La Jolla, CA, USA) according to the manufacturer's instructions. An *E. coli* transformant containing a plasmid of the correct size was detected by restriction digestion and plasmid DNA was prepared using a BIOROBOT® 9600 (QIAGEN Inc., Valencia, CA, USA). DNA sequencing of the insert from this plasmid was performed with an Applied Biosystems Model 377 XL Automated DNA Sequencer (Perkin-Elmer/Applied Biosystems, Inc., Foster City, CA, USA) using dye-terminator chemistry (Giesecke et al., 1992, Journal of Virology Methods 38: 47-60) and primer walking strategy. Nucleotide sequence data were scrutinized for quality

and all sequences were compared to each other with assistance of PHRED/PHRAP software (University of Washington, Seattle, WA, USA). The nucleotide sequence was shown to match the genomic sequence determined by TIGR. The resulting plasmid was designated pEJG93.

[0393] *Aspergillus oryzae* JaL250 (WO 99/61651) protoplasts were prepared according to the method of Christensen et al., 1988, Bio/Technology 6: 1419-1422, and transformed with 5 µg of pEJG93 (as well as pAIIo2 as a vector control). The transformation yielded about 100 transformants. Ten transformants were isolated to individual PDA plates.

[0394] Confluent PDA plates of five of the ten transformants were washed with 5 ml of 0.01% TWEEN® 20 and inoculated separately into 25 ml of MDU2BP medium in 125 ml glass shake flasks and incubated at 34°C, 250 rpm. Five days after incubation, 0.5 µl of supernatant from each culture was analyzed using 8-16% Tris-Glycine SDS-PAGE gels (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. SDS-PAGE profiles of the cultures showed that one of the transformants had a major band of approximately 70 kDa. This transformant was named *Aspergillus oryzae* JaL250EJG93.

[0395] One hundred ml of shake flask medium were added to a 500 ml shake flask. The shake flask medium was composed of 50 g of sucrose, 10 g of KH₂PO₄, 0.5 g of CaCl₂, 2 g of MgSO₄·7H₂O, 2 g of K₂SO₄, 2 g of urea, 10 g of yeast extract, 2 g of citric acid, 0.5 ml of trace metals solution, and deionized water to 1 liter. The trace metals solution was composed of 13.8 g of FeSO₄·7H₂O, 14.3 g of ZnSO₄·7H₂O, 8.5 g of MnSO₄·H₂O, 2.5 g of CuSO₄·5H₂O, 3 g of citric acid, and deionized water to 1 liter. The shake flask was inoculated with two plugs of *Aspergillus oryzae* JaL250EJG93 from a PDA plate and incubated at 34°C on an orbital shaker at 200 rpm for 24 hours. Fifty ml of the shake flask broth were used to inoculate a 3 liter fermentation vessel.

[0396] A total of 1.8 liters of the fermentation batch medium was added to a three liter glass jacketed fermentor (Applikon Biotechnology, Schiedam, Netherlands). The fermentation batch medium was composed per liter of 10 g of yeast extract, 24 g of sucrose, 5 g of (NH₄)₂SO₄, 2 g of KH₂PO₄, 0.5 g of CaCl₂·2H₂O, 2 g of MgSO₄·7H₂O, 1 g of citric acid, 2 g of K₂SO₄, 0.5 ml of anti-foam, and 0.5 ml of trace metals solution. The trace metals solution was composed of 13.8 g of FeSO₄·7H₂O, 14.3 g of ZnSO₄·7H₂O, 8.5 g of MnSO₄·H₂O, 2.5 g of CuSO₄·5H₂O, 3 g of citric acid, and deionized water to 1 liter. Fermentation feed medium was composed of maltose. The fermentation feed medium was dosed at a rate of 0 to 4.4 g/l/hr for a period of 185 hours. The fermentation vessel was maintained at a temperature of 34°C and pH was controlled using an Applikon 1030 control system (Applikon Biotechnology, Schiedam, Netherlands) to a set-point of 6.1 +/- 0.1. Air was added to the vessel at a rate of 1 vvm and the broth was agitated by a Rushton impeller rotating at 1100 to 1300 rpm. At the end of the fermentation, whole broth was harvested from the vessel and centrifuged at 3000 x g to remove the biomass. The supernatant was sterile filtered and stored at 5 to 10°C.

[0397] The filtered broth was concentrated and buffer exchanged with 20 mM Tris-HCl pH 8.5

using a tangential flow concentrator (Pall Filtron, Northborough, MA, USA) equipped with a 10 kDa polyethersulfone membrane (Pall Filtron, Northborough, MA, USA). The sample was loaded onto a Q SEPHAROSE® High Performance column (GE Healthcare, Piscataway, NJ, USA) equilibrated in 20 mM Tris pH 8.0, and bound proteins were eluted with a linear gradient from 0-600 mM sodium chloride. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit (Thermo Fischer Scientific, Waltham, MA, USA) in which bovine serum albumin was used as a protein standard.

Example 6: Preparation of *Aspergillus fumigatus* cellobiohydrolase II

[0398] *Aspergillus fumigatus* NN055679 GH6A cellobiohydrolase II (CBHII) (SEQ ID NO: 9 [DNA sequence] and SEQ ID NO: 10 [deduced amino acid sequence]) was prepared according to the following procedure.

[0399] Two synthetic oligonucleotide primers, shown below, were designed to PCR amplify the full-length open reading frame of the *Aspergillus fumigatus* cellobiohydrolase II gene from genomic DNA. A TOPO® Cloning Kit (Invitrogen Corp., Carlsbad, CA, USA) was used to clone the PCR product. An IN-FUSION™ Cloning Kit was used to clone the fragment into pAILo2.

Forward primer:

5'-ACTGGATTTACCAT**GAAGCACCTTGCATCTTCCATCG**-3' (SEQ ID NO: 11)

Reverse primer:

5'-TCACCTCTAGTTAATTA**AAAGGACGGGTTAGCGT**-3' (SEQ ID NO: 12)

Bold letters represent coding sequence. The remaining sequence contains sequence identity compared with the insertion sites of pAILo2.

[0400] Fifty picomoles of each of the primers above were used in a PCR reaction containing 500 ng of *Aspergillus fumigatus* genomic DNA, 1X ThermoPol *Taq* reaction buffer (New England Biolabs, Ipswich, MA, USA), 6 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, and 0.1 unit of *Taq* DNA polymerase (New England Biolabs, Ipswich, MA, USA), in a final volume of 50 µl. An EPPENDORF® MASTERCYCLER® 5333 (Eppendorf Scientific, Inc., Westbury, NY, USA) was used to amplify the fragment programmed for 1 cycle at 98°C for 2 minutes; and 35 cycles each at 96°C for 30 seconds, 61°C for 30 seconds, and 72°C for 2 minutes. After the 35 cycles, the reaction was incubated at 72°C for 10 minutes and then cooled at 10°C until further processed. To remove the A-tails produced by the *Taq* DNA polymerase the reaction was incubated for 10 minutes at 68°C in the presence of 1 unit of *Pfx* DNA polymerase (Invitrogen, Carlsbad, CA, USA).

[0401] A 1.3 kb PCR reaction product was isolated by 0.8% GTG-agarose gel electrophoresis (Cambrex Bioproducts, East Rutherford, NJ, USA) using TAE buffer and 0.1 µg of ethidium bromide per ml. The DNA band was visualized with the aid of a DARKREADER™ Transilluminator to avoid UV-induced mutations. The 1.3 kb DNA band was excised with a

disposable razor blade and purified using an ULTRAFREE®-DA spin cup (Millipore, Billerica, MA) according to the manufacturer's instructions.

[0402] The purified 1.3 kb PCR product was cloned into vector pCR®4Blunt-TOPO® (Invitrogen, Carlsbad, CA, USA). Two µl of the purified PCR product were mixed with 1 µl of a 2 M sodium chloride solution and 1 µl of the vector. The reaction was incubated at room temperature for 15 minutes and then 2 µl of the reaction were transformed into *E. coli* TOP10 competent cells (Invitrogen Corp., Carlsbad, CA, USA) according to the manufacturer's instructions. Two aliquots of 100 µl each of the transformation reaction were spread onto two 150 mm 2XYT plates supplemented with 100 µg of ampicillin per ml and incubated overnight at 37°C.

[0403] Eight recombinant colonies were used to inoculate liquid cultures containing 3 ml of LB medium supplemented with 100 µg of ampicillin per ml. Plasmid DNA was prepared from these cultures using a BIOROBOT® 9600. Clones were analyzed by restriction digestion. Plasmid DNA from each clone was digested with *Eco* RI and analyzed by agarose gel electrophoresis as above. Six out of eight clones had the expected restriction digestion pattern and from these, clones 2, 4, 5, 6, 7 and 8 were selected to be sequenced to confirm that there were no mutations in the cloned insert. Sequence analysis of their 5-prime and 3-prime ends indicated that clones 2, 6 and 7 had the correct sequence. These three clones were selected for re-cloning into pAIlO2. One microliter aliquot of each clone was mixed with 17 µl of 10-fold diluted 0.1 M EDTA-10 mM Tris (TE) and 1 µl of this mix was used to re-amplify the *Aspergillus fumigatus* GH6A coding region.

[0404] Fifty picomoles of each of the primers above were used in a PCR reaction containing 1 µl of the diluted mix of clones 2, 6 and 7, 1X *Pfx* Amplification Buffer (Invitrogen, Carlsbad, CA, USA), 6 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 2.5 units of PLATINUM® *Pfx* DNA polymerase (Invitrogen, Carlsbad, CA, USA), and 1 µl of 50 mM MgSO₄, in a final volume of 50 µl. An EPPENDORF® MASTERCYCLER® 5333 was used to amplify the fragment programmed for 1 cycle at 98°C for 2 minutes; and 35 cycles each at 94°C for 30 seconds, 61°C for 30 seconds, and 68°C for 1.5 minutes. After the 35 cycles, the reaction was incubated at 68°C for 10 minutes and then cooled to 10°C until further processed. A 1.3 kb PCR reaction product was isolated by 0.8% GTG-agarose gel electrophoresis using TAE buffer and 0.1 µg of ethidium bromide per ml. The DNA band was visualized with the aid of a DARKREADER™ Transilluminator to avoid UV-induced mutations. The 1.0 kb DNA band was excised with a disposable razor blade and purified with an ULTRAFREE®-DA spin cup according to the manufacturer's instructions.

[0405] The vector pAIlO2 was linearized by digestion with *Nco* I and *Pac* I. The fragment was purified by gel electrophoresis and ultrafiltration as described above. Cloning of the purified PCR fragment into the linearized and purified pAIlO2 vector was performed with an IN-FUSION™ Cloning Kit. The reaction (20 µl) contained 2 µl of 1X IN-FUSION™ Buffer, 1X BSA, 1 µl of IN-FUSION™ enzyme (diluted 1:10), 100 ng of pAIlO2 digested with *Nco* I and *Pac* I, and 50 ng of the *Aspergillus fumigatus* GH6A purified PCR product. The reaction was

incubated at room temperature for 30 minutes. A 2 µl sample of the reaction was used to transform *E. coli* TOP10 competent cells according to the manufacturer's instructions. After the recovery period, two 100 µl aliquots from the transformation reaction were plated onto 150 mm 2XYT plates supplemented with 100 µg of ampicillin per ml. The plates were incubated overnight at 37°C. A set of eight putative recombinant clones was selected at random from the selection plates and plasmid DNA was prepared from each one using a BIOROBOT® 9600. Clones were analyzed by *Pst* I restriction digest. Seven out of eight clones had the expected restriction digestion pattern. Clones 1, 2, and 3 were then sequenced to confirm that there were no mutations in the cloned insert. Clone #2 was selected and designated pAlLo33.

[0406] *Aspergillus oryzae* JaL355 protoplasts were prepared according to the method of Christensen et al., 1988, Bio/Technology 6: 1419-1422, which were transformed with 5 µg of plasmid pAlLo33. The transformation yielded about 30 transformants. Twenty-six transformants were isolated to individual PDA plates.

[0407] Confluent PDA plates of four of the transformants were washed with 8 ml of 0.01% TWEEN® 20 and inoculated separately into 1 ml of MDU2BP medium in sterile 24 well tissue culture plates and incubated at 34°C. Three days after incubation, 20 µl of harvested broth from each culture were analyzed using 8-16% Tris-Glycine SDS-PAGE gels (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. SDS-PAGE profiles of the cultures showed that several transformants had a new major band of approximately 75 kDa. One transformant was designated *Aspergillus oryzae* JaL355 ALLO33 (EXP03191).

[0408] One hundred ml of shake flask medium were added to a 500 ml shake flask. The shake flask medium was composed of 50 g of sucrose, 10 g of KH₂PO₄, 0.5 g of CaCl₂, 2 g of MgSO₄·7H₂O, 2 g of K₂SO₄, 2 g of urea, 10 g of yeast extract, 2 g of citric acid, 0.5 ml of trace metals solution, and deionized water to 1 liter. The trace metals solution was composed of 13.8 g of FeSO₄·7H₂O, 14.3 g of ZnSO₄·7H₂O, 8.5 g of MnSO₄·H₂O, 2.5 g of CuSO₄·5H₂O, 3 g of citric acid, and deionized water to 1 liter. The shake flask was inoculated with two plugs of *Aspergillus oryzae* JaL355 ALLO33 (EXP03191) from a PDA plate and incubated at 34°C on an orbital shaker at 200 rpm for 24 hours. Fifty ml of the shake flask broth were used to inoculate a 3 liter fermentation vessel.

[0409] A total of 1.8 liters of the fermentation batch medium was added to a three liter glass jacketed fermentor. The fermentation batch medium was composed per liter of 10 g of yeast extract, 24 g of sucrose, 5 g of (NH₄)₂SO₄, 2 g of KH₂PO₄, 0.5 g of CaCl₂·2H₂O, 2 g of MgSO₄·7H₂O, 1 g of citric acid, 2 g of K₂SO₄, 0.5 ml of anti-foam, and 0.5 ml of trace metals solution. The trace metals solution was composed of 13.8 g of FeSO₄·7H₂O, 14.3 g of ZnSO₄·7H₂O, 8.5 g of MnSO₄·H₂O, 2.5 g of CuSO₄·5H₂O, 3 g of citric acid, and deionized water to 1 liter. Fermentation feed medium was composed of maltose. The fermentation feed medium was dosed at a rate of 0 to 4.4 g/l/hr for a period of 185 hours. The fermentation vessel was maintained at a temperature of 34°C and pH was controlled using an Applikon 1030 control system to a set-point of 6.1 +/- 0.1. Air was added to the vessel at a rate of 1 vvm

and the broth was agitated by a Rushton impeller rotating at 1100 to 1300 rpm. At the end of the fermentation, whole broth was harvested from the vessel and centrifuged at 3000 x *g* to remove the biomass. The supernatant was sterile filtered and stored at 5-10°C. The broth was filtered using a 0.22 µm EXPRESS™ Plus Membrane (Millipore, Bedford, MA, USA).

[0410] A 100 ml volume of the filtered broth was buffer exchanged into 20 mM Tris pH 8.0 using a 400 ml SEPHADEX™ G-25 column (GE Healthcare, United Kingdom) according to the manufacturer's instructions. The fractions were pooled and adjusted to 1.2 M ammonium sulphate-20 mM Tris pH 8.0. The equilibrated protein was loaded onto a PHENYL SEPHAROSE™ 6 Fast Flow (high sub) column equilibrated in 20 mM Tris pH 8.0 with 1.2 M ammonium sulphate, and bound proteins were eluted with 20 mM Tris pH 8.0 with no ammonium sulphate. The fractions were pooled and protein concentration was determined using a Microplate BCA™ Protein Assay Kit in which bovine serum albumin was used as a protein standard.

Example 7: Preparation of *Thermoascus aurantiacus* CGMCC 0583 GH61A polypeptide having cellulolytic enhancing activity

[0411] *Thermoascus aurantiacus* CGMCC 0583 GH61A polypeptide having cellulolytic enhancing activity (SEQ ID NO: 13 [DNA sequence] and SEQ ID NO: 14 [deduced amino acid sequence]) was recombinantly prepared according to WO 2005/074656 using *Aspergillus oryzae* JaL250 as a host. The recombinantly produced *Thermoascus aurantiacus* GH61A polypeptide was first concentrated by ultrafiltration using a 10 kDa membrane, buffer exchanged into 20 mM Tris-HCl pH 8.0, and then purified using 320 ml SUPERDEX® 75 SEC column (GE Healthcare, Piscataway, NJ, USA) with isocratic elution of approximately 1.3 liters of 150 mM NaCl-20 mM Tris pH 8.0. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit in which bovine serum albumin was used as a protein standard.

Example 8: Preparation of *Aspergillus fumigatus* NN055679 Cel3A beta-glucosidase

[0412] *Aspergillus fumigatus* NN055679 Cel3A beta-glucosidase (SEQ ID NO: 15 [DNA sequence] and SEQ ID NO: 16 [deduced amino acid sequence]) was recombinantly prepared according to WO 2005/047499 using *Trichoderma reesei* RutC30 as a host.

[0413] The broth was filtered using a 0.22 µm EXPRESS™ Plus Membrane. The filtered broth was concentrated and buffer exchanged into 20 mM Tris-HCl pH 8.5 using a tangential flow concentrator equipped with a 10 kDa polyethersulfone membrane. The sample was loaded onto a Q SEPHAROSE® High Performance column equilibrated in 20 mM Tris pH 8.0, and bound proteins were eluted with a linear gradient from 0-600 mM sodium chloride. The fractions were concentrated and loaded onto a SUPERDEX® 75 HR 26/60 column (GE Healthcare, Piscataway, NJ, USA) equilibrated with 20 mM Tris-150 mM sodium chloride pH

8.5. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit in which bovine serum albumin was used as a protein standard.

Example 9: Preparation of *Aspergillus fumigatus* NN055679 GH10 xylanase

[0414] *Aspergillus fumigatus* NN055679 GH10 xylanase (xyn3) (SEQ ID NO: 17 [DNA sequence] and SEQ ID NO: 18 [deduced amino acid sequence]) was prepared recombinantly according to WO 2006/078256 using *Aspergillus oryzae* BECh2 (WO 2000/039322) as a host.

[0415] The broth was filtered using a 0.22 µm EXPRESS™ Plus Membrane. A 100 ml volume of filtered broth was buffer exchanged into 50 mM sodium acetate pH 5.0 using a 400 ml SEPHADEX™ G-25 column according to the manufacturer's instructions. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit with bovine serum albumin as a protein standard.

Example 10: Preparation of *Talaromyces emersonii* CBS 393.64 GH3 beta-xylosidase

[0416] *Talaromyces emersonii* CBS 393.64 (NN005049) beta-xylosidase (SEQ ID NO: 19 [DNA sequence] and SEQ ID NO: 20 [deduced amino acid sequence]) was prepared recombinantly according to Rasmussen et al., 2006, Biotechnology and Bioengineering 94: 869-876 using *Aspergillus oryzae* JaL355 (WO 2003/070956) as a host.

[0417] The broth was filtered using a 0.22 µm EXPRESS™ Plus Membrane. A 100 ml volume of filtered broth was buffer exchanged into 50 mM sodium acetate pH 5.0 using a 400 ml SEPHADEX™ G-25 column according to the manufacturer's instructions. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit with bovine serum albumin as a protein standard.

Example 11: Effect of *Chaetomium virescens* GH5 endoglucanase on a high-temperature enzyme composition using milled unwashed PCS at 50-65°C

[0418] The *Chaetomium virescens* GH5 endoglucanase (EXP03745) was evaluated in a high-temperature enzyme composition at 50°C, 55°C, 60°C, and 65°C using milled unwashed PCS as a substrate. The high-temperature enzyme composition included 37% *Aspergillus fumigatus* Cel7A cellobiohydrolase I, 25% *Aspergillus fumigatus* Cel6A cellobiohydrolase II, 10% *Chaetomium virescens* Family GH5 endoglucanase II, 15% *Thermoascus aurantiacus* GH61A polypeptide having cellulolytic enhancing activity, 5% *Aspergillus fumigatus* Cel3A beta-glucosidase, 5% *Aspergillus fumigatus* GH10 xyn3 xylanase, and 3% *Talaromyces emersonii* GH3 beta-xylosidase. The high-temperature enzyme composition was added to PCS hydrolysis reactions at 3.5 mg total protein per g cellulose, and the hydrolysis results were compared with

the results for a similar high-temperature enzyme composition without the *C. virescens* GH5 endoglucanase (3.15 mg protein per g cellulose).

[0419] The assay was performed as described in Example 4. The 1 ml reactions with milled unwashed PCS (5% insoluble solids) were conducted for 72 hours in 50 mM sodium acetate pH 5.0 buffer containing 1 mM manganese sulfate. All reactions were performed in triplicate and involved single mixing at the beginning of hydrolysis.

[0420] The results shown in Figure 2 demonstrated that at 50°C, 55°C, 60°C, and 65°C the high-temperature enzyme composition that included 10% *C. virescens* GH5 endoglucanase significantly outperformed the enzyme composition containing no *C. virescens* GH5 endoglucanase.

Example 12: Effect of *Chaetomium virescens* GH5 endoglucanase on unbleached eucalyptus Kraft pulp

[0421] The *Chaetomium virescens* GH5 endoglucanase was assessed for its ability to improve the strength properties of paper made from eucalyptus Kraft pulp.

[0422] Brazilian unbleached eucalyptus Kraft pulp was treated with the *C. virescens* GH5 endoglucanase in a 1000 ml Lab-O-Mat beaker (Werner Mathis AG, Zurich, Switzerland) at 4% consistency, pH 6 and 50°C for 1 hour. The amount of pulp was 15 g and the enzyme dosage was 6 mg of *C. virescens* GH5 endoglucanase per kg dry pulp. After the enzyme treatment, the sample was diluted to 2 liter with deionized water and the slurry was disintegrated for 15,000 revolutions in a standard Pulp Disintegrator (Type 8-3; Lorentzen & Wettre, Kista, Sweden). The reference pulp (negative control) was treated in the same way but without enzyme addition. Handsheets were prepared according to TAPPI test method, "Forming handsheets for physical testing of pulp" T-205 sp-95. For the determination of tensile index (Instron Model 5564) and tear index (Digital Elemendorf Tear Tester) tests were conducted according to TAPPI Test Methods T494 om-96 and T414 om-98, respectively.

[0423] Figures 3 and 4 present the physical data obtained from handsheets made from *C. virescens* GH5 endoglucanase treatment of unbleached eucalyptus Kraft pulp. As shown in Figure 3, treatment of the unbleached eucalyptus Kraft pulp with the *C. virescens* GH5 endoglucanase improved the handsheet tensile index by 33% relative to the control. The *C. virescens* GH5 endoglucanase treatment resulted in approximately 17% increase in handsheet tear index (Figure 4).

Example 13: Preparation of *Trichoderma reesei* RutC30 GH3 beta-xylosidase

[0424] A *Trichoderma reesei* RutC30 GH3 beta-xylosidase gene (SEQ ID NO: 31 [DNA

sequence] and SEQ ID NO: 32) was prepared in *Trichoderma reesei* and purified according to WO 2011/057140.

Example 14: Preparation of *Trichoderma reesei* GH5 endoglucanase II

[0425] The *Trichoderma reesei* GH5 endoglucanase II (SEQ ID NO: 33 [DNA sequence] and SEQ ID NO: 34 [deduced amino acid sequence]) was prepared recombinantly according to WO 2011/057140 using *Aspergillus oryzae* as a host. The filtered broth was concentrated and buffer exchanged using a tangential flow concentrator equipped with a 10 kDa polyethersulfone membrane with 20 mM Tris-HCl pH 8.0. The buffer-exchanged protein was loaded onto a 20 ml MONO Q® column (GE Healthcare, Piscataway, NJ, USA) equilibrated in 20 mM Tris pH 8.0, and bound proteins were eluted with a linear gradient from 0-600 mM sodium chloride. The fractions were pooled based on SDS-PAGE analysis using a 8-16% Tris HCl CRITERION® Stain Free gels (Bio-Rad Laboratories, Inc., Hercules, CA, USA) with Precision Plus Protein™ Unstained Standards (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Protein concentration was determined using a Microplate BCA™ Protein Assay Kit in which bovine serum albumin was used as a protein standard.

Example 15: Comparison of the effect of *Chaetomium virescens* GH5 endoglucanase and *Trichoderma reesei* GH5 endoglucanase II in the hydrolysis of milled unwashed PCS by a cellulase enzyme composition

[0426] *Chaetomium virescens* GH5 endoglucanase and *Trichoderma reesei* GH5 endoglucanase II were evaluated in a cellulase enzyme composition at 50°C and 55°C using milled unwashed PCS as a substrate. The cellulase enzyme composition included 37% *Aspergillus fumigatus* Cel7A cellobiohydrolase I, 25% *Aspergillus fumigatus* Cel6A cellobiohydrolase II, 10% endoglucanase, 15% *Thermoascus aurantiacus* GH61A polypeptide having cellulolytic enhancing activity, 5% *Aspergillus fumigatus* Cel3A beta-glucosidase, 5% *Aspergillus fumigatus* GH10 xyn3 xylanase, and 3% *Trichoderma reesei* GH3 beta-xylosidase. Each endoglucanase was added individually at 0.35 mg enzyme protein per g cellulose to 3.15 mg enzyme protein of the cellulase enzyme composition without endoglucanase per g cellulose.

[0427] The assay was performed as described in Example 4. The 1 ml reactions with milled unwashed PCS (5% insoluble solids) were conducted for 72 hours in 50 mM sodium acetate pH 5.0 buffer containing 1 mM manganese sulfate. All reactions were performed in triplicate and involved single mixing at the beginning of hydrolysis.

[0428] The results shown in Figure 5 demonstrated that at 50°C and 55°C, the cellulase enzyme composition that included 10% *Chaetomium virescens* GH5 endoglucanase had significantly higher glucose conversion than the cellulase enzyme composition with 10%

Trichoderma reesei GH5 endoglucanase II.

Example 16: Specific activity of *Chaetomium virescens* GH5 endoglucanase on carboxymethyl cellulose

[0429] The specific activity of the *Chaetomium virescens* GH5 endoglucanase was determined using carboxymethyl cellulose (CMC; Hercules Chemical Company, Inc., Passaic, NJ, USA) as substrate at 10 g per liter of 50 mM sodium acetate pH 5.0. To 190 μ l of the CMC solution, 10 μ l of *Chaetomium virescens* GH5 endoglucanase (at different loadings) were added. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit in which bovine serum albumin was used as a protein standard. Substrate control and enzyme control were included. The reaction was incubated at 50°C for 30 minutes followed by addition of 50 μ l of 0.5 M sodium hydroxide to stop the reaction. The reducing sugars produced were determined using a para-hydroxybenzoic acid hydrazide (PHBAH, Sigma, St. Louis, MO) assay adapted to a 96 well microplate format as described below. Briefly, a 100 μ l aliquot of an appropriately diluted sample was placed in a 96-well conical bottomed microplate. Reactions were initiated by adding 50 μ l of 1.5% (w/v) PHBAH in 2% sodium hydroxide to each well. Plates were heated uncovered at 95°C for 10 minutes. Plates were allowed to cool to room temperature and 50 μ l of deionized water were added to each well. A 100 μ l aliquot from each well was transferred to a flat bottomed 96 well plate and the absorbance at A_{410nm} measured using a SPECTRAMAX® Microplate Reader (Molecular Devices, Sunnyvale, CA). Glucose standards (0.1-0.0125 mg/ml diluted with 0.4% sodium hydroxide) were used to prepare a standard curve to translate the obtained A_{410nm} values into glucose equivalents. The enzyme loading versus the reducing sugars produced was plotted and the linear range was used to calculate the specific activity of *Chaetomium virescens* GH5 endoglucanase, as expressed as μ mol glucose equivalent produced per minute per mg enzyme, or IU/mg.

[0430] The results indicated that the specific activity of the *Chaetomium virescens* GH5 endoglucanase on CMC was 25.1 IU/mg protein.

SEQUENCE LISTING

[0431]

<110> Novozymes, Inc. Novozymes A/S Spodsborg, Nikolaj McBrayer, Brett Cassland, Pierre

<120> Polypeptides having endoglucanase activity and polynucleotides encoding same

<130> 12167-WO-PCT

<150> US 61/480,245

<151> 2011-04-28

<160> 34

<170> PatentIn version 3.5

<210> 1

<211> 1320

<212> DNA

<213> Chaetophractus villosus

<400> 1

atgccttcag	tcgtcctttc	tcttgctttc	ttggcgggca	gtgcccttgc	ccagcagagt	60
gtctggggac	aatgcggtgg	aataggctgg	attgggccaa	caaactgcgc	ggctgggagc	120
tgctgcgcta	cccagaacgc	ctggtatgca	caatgtacgc	ccggggtggc	atgcggtggt	180
ggtggtagtg	gtacgacttt	gacaacctg	acaacccgga	ccacgactgt	tcaggccacg	240
actacgaccg	ttacatcgac	ccccaccagc	aacggcaagg	ttcgtctcgc	cggcgtaaac	300
attgcaggct	tcgactttgg	gtgcgggacc	gatggcaact	gtaataccgc	aaaggtgtgg	360
ccgccagtga	agaactaccc	cccagactac	aaccaccccg	acggcgccgg	tcagatgcag	420
cacttctaca	aggatgacca	catgaacatc	ttccgcctgc	ccgtgggttg	gcaatacctg	480
gtcaacaaca	acctcggcgg	caacttgac	agcaccaact	tcggctacta	cgaccagctg	540
gtccaggcct	gcctcaacac	gggcgcatac	tgcatcatcg	acatccacaa	ctacgcgcgg	600
tggaatgggg	ccattatcgg	ccagggcggg	cccacaaacg	agcagttcgt	cagcatatgg	660
aagcagctgg	cgagcaagta	gcgccgccag	tcgcgcatac	ggttcggcat	catgaacgag	720
ccccacgaca	tacccaatat	caacacctgg	gcggccacgg	ttcaggccgt	cgtcacagcc	780
atccgcaatg	ccggcgccac	cagccagttc	atctcgctgc	ccggcaacgg	gtggcagtc	840
gcgggcacct	tcatcagcga	cggaagcgct	gcggctctgt	ctaccgtcag	gaaccgggat	900
ggatcgacga	cgaacctcat	ttttgatgtg	cacaagtacc	tcgattctga	caactcgggg	960
acgcataccg	agtgtgtcaa	gaataatatc	gacgacgcgt	ttgcgccgct	tgccgcctgg	1020
ctcaggcaaa	ataacagaca	ggcgatcctc	actgagacgg	gaggcggcaa	cacggcttca	1080
tgccagactt	atgtgtgcc	gcagattgcg	tacctgaagt	aagtattatc	gacatcgttt	1140
ttcctttgta	ggacaacttg	ctcaccttca	tgcttttagtc	aaaactcgga	cgtctatctc	1200
ggctatgtcg	gctgggctgc	tggttcgttc	gacaatacgt	acgagctggt	cgagacgccg	1260
acctggaatg	gtaactcggt	tagtgatacg	gcccttgta	ggtcttgcc	tgctcgtaa	1320

<210> 2

<211> 419

<212> PRT

<213> Chaetophractus villosus

<400> 2

Met	Pro	Ser	Val	Val	Leu	Ser	Leu	Ala	Phe	Leu	Ala	Gly	Ser	Ala	Leu
1				5					10					15	

Ala	Gln	Gln	Ser	Val	Thr	Gly	Gln	Gln	Gln	Gln	Gln	Gln	Thr	Thr	Gln
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

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

Thr His Thr Glu Cys Val Lys Asn Asn Ile Asp Asp Ala Phe Ala Pro
325 330 335

Leu Ala Ala Trp Leu Arg Gln Asn Asn Arg Gln Ala Ile Leu Thr Glu
340 345 350

Thr Gly Gly Gly Asn Thr Ala Ser Cys Gln Thr Tyr Val Cys Gln Gln
355 360 365

Ile Ala Tyr Leu Asn Gln Asn Ser Asp Val Tyr Leu Gly Tyr Val Gly
370 375 380

Trp Ala Ala Gly Ser Phe Asp Asn Thr Tyr Glu Leu Val Glu Thr Pro
385 390 395 400

Thr Trp Asn Gly Asn Ser Phe Ser Asp Thr Ala Leu Val Arg Ser Cys
405 410 415

Leu Ala Arg

<210> 3

<211> 41

<212> DNA

<213> Chaetophractus villosus

<400> 3

acacaactgg ggatccacca tgccttcagt cgtcctttct c 41

<210> 4

<211> 37

<212> DNA

<213> Chaetophractus villosus

<400> 4

ccctctagat ctcgagcaac aatgccgaca cactcca 37

<210> 5

<211> 1599

<212> DNA

<213> Aspergillus fumigatus

<400> 5

atgctggcct ccaccttctc ctaccgcatg tacaagaccg cgctcatcct ggccgccctt 60

ctgggctctg gccaggctca gcaggctcgt acttcccagg cggaagtgca tccgtccatg 120

acctggcaga gctgcacggc tggcggcagc tgcaccacca acaacggcaa ggtggtcatc 180

gacgcgaact ggcgttgggt gcacaaagtc ggcgactaca ccaactgcta caccggcaac 240

acctgggaca cgactatctg ccttgacgat gcgacctgog catccaactg cgcccttgag 300

ggtgcccaact acgaatccac ctatggtgtg accgccagcg gcaattccct ccgcctcaac 360

```

ttcgtcacca ccagccagca gaagaacatt ggctcgcgtc tgtacatgat gaaggacgac      420
tcgacctacg agatgtttaa gctgctgaac caggagtcca ccttcgatgt cgatgtctcc      480
aacctcccct gcggtctcaa cggtgctctg tactttgtcg ccatggacgc cgacggtggc      540
atgtccaagt acccaaccaa caaggccggt gccaaagtacg gtactggata ctgtgactcg      600
cagtgccttc gcgacctcaa gttcatcaac ggtcaggcca acgtcgaagg gtggcagccc      660
tcctccaacg atgccaatgc gggtagcggc aaccacgggt cctgctgcgc ggagatggat      720
atctgggagg ccaacagcat ctccacggcc ttcaccccc atccgtgcga cacgcccggc      780
caggtgatgt gcaccggtga tgccctgcgt ggcacctaca gctccgaccg ctacggcggc      840
acctgcgacc ccgacggatg tgatttcaac tccttccgcc agggcaacaa gaccttctac      900
ggccttgcca tgaccgtcga caccaagagc aagtttacgg tcgtcaccca gttcatcacc      960
gacgacggca cctccagcgg caccctcaag gagatcaagc gcttctacgt gcagaacggc     1020
aaggtgatcc ccaactcgga gtgcacctgg accggcgtca gcggcaactc catcaccacc     1080
gagtactgca ccgcccagaa gagcctgttc caggaccaga acgtcttcga aaagcacggc     1140
ggcctcgagg gcatgggtgc tgccctcgcc cagggtatgg ttctcgtcac gtccctgtgg     1200
gatgatcact cggccaacat gctctggttc gacagcaact acccgaccac tgccctctcc     1260

accactcccg gcgtcgcccc tggtagctgc gacatctcct ccggcgctcc tgccgatgtc     1320
gaggcgaacc accccgacgc ctacgtcgtc tactccaaca tcaaggtcgg ccccatcggc     1380
tcgaccttca acagcgggtg ctogaacccc ggtggcggaa ccaccacgac aactaccacc     1440
cagcctacta ccaccacgac cacggctgga aaccctggcg gcaccggagt cgcacagcac     1500
tatggccagt gtggtggaat cggatggacc ggaccacaaa cctgtgccag cccttatacc     1560
tgccagaagc tgaatgatta ttactctcag tgccctgtag                               1599

```

<210> 6

<211> 532

<212> PRT

<213> *Aspergillus fumigatus*

<400> 6

```

Met Leu Ala Ser Thr Phe Ser Tyr Arg Met Tyr Lys Thr Ala Leu Ile
 1              5              10              15

Leu Ala Ala Leu Leu Gly Ser Gly Gln Ala Gln Gln Val Gly Thr Ser
 20              25              30

Gln Ala Glu Val His Pro Ser Met Thr Trp Gln Ser Cys Thr Ala Gly
 35              40              45

Gly Ser Cys Thr Thr Asn Asn Gly Lys Val Val Ile Asp Ala Asn Trp
 50              55              60

Arg Trp Val His Lys Val Gly Asp Tyr Thr Asn Cys Tyr Thr Gly Asn
 65              70              75              80

Thr Trp Asp Thr Thr Ile Cys Pro Asp Asp Ala Thr Cys Ala Ser Asn
--              --              --

```

85	90	95
Cys Ala Leu Glu Gly Ala Asn Tyr Glu Ser Thr Tyr Gly Val Thr Ala		
100	105	110
Ser Gly Asn Ser Leu Arg Leu Asn Phe Val Thr Thr Ser Gln Gln Lys		
115	120	125
Asn Ile Gly Ser Arg Leu Tyr Met Met Lys Asp Asp Ser Thr Tyr Glu		
130	135	140
Met Phe Lys Leu Leu Asn Gln Glu Phe Thr Phe Asp Val Asp Val Ser		
145	150	155
Asn Leu Pro Cys Gly Leu Asn Gly Ala Leu Tyr Phe Val Ala Met Asp		
165	170	175
Ala Asp Gly Gly Met Ser Lys Tyr Pro Thr Asn Lys Ala Gly Ala Lys		
180	185	190
Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys Pro Arg Asp Leu Lys Phe		
195	200	205
Ile Asn Gly Gln Ala Asn Val Glu Gly Trp Gln Pro Ser Ser Asn Asp		
210	215	220
Ala Asn Ala Gly Thr Gly Asn His Gly Ser Cys Cys Ala Glu Met Asp		
225	230	235
Ile Trp Glu Ala Asn Ser Ile Ser Thr Ala Phe Thr Pro His Pro Cys		
245	250	255
Asp Thr Pro Gly Gln Val Met Cys Thr Gly Asp Ala Cys Gly Gly Thr		
260	265	270
Tyr Ser Ser Asp Arg Tyr Gly Gly Thr Cys Asp Pro Asp Gly Cys Asp		
275	280	285
Phe Asn Ser Phe Arg Gln Gly Asn Lys Thr Phe Tyr Gly Pro Gly Met		
290	295	300
Thr Val Asp Thr Lys Ser Lys Phe Thr Val Val Thr Gln Phe Ile Thr		
305	310	315
Asp Asp Gly Thr Ser Ser Gly Thr Leu Lys Glu Ile Lys Arg Phe Tyr		
325	330	335
Val Gln Asn Gly Lys Val Ile Pro Asn Ser Glu Ser Thr Trp Thr Gly		
340	345	350
Val Ser Gly Asn Ser Ile Thr Thr Glu Tyr Cys Thr Ala Gln Lys Ser		
355	360	365
Leu Phe Gln Asp Gln Asn Val Phe Glu Lys His Gly Gly Leu Glu Gly		
370	375	380
Met Gly Ala Ala Leu Ala Gln Gly Met Val Leu Val Met Ser Leu Trp		

```

385                               390                               395                               400

Asp Asp His Ser Ala Asn Met Leu Trp Leu Asp Ser Asn Tyr Pro Thr
      405                          410                          415

Thr Ala Ser Ser Thr Thr Pro Gly Val Ala Arg Gly Thr Cys Asp Ile
      420                          425                          430

Ser Ser Gly Val Pro Ala Asp Val Glu Ala Asn His Pro Asp Ala Tyr
      435                          440                          445

Val Val Tyr Ser Asn Ile Lys Val Gly Pro Ile Gly Ser Thr Phe Asn
      450                          455                          460

Ser Gly Gly Ser Asn Pro Gly Gly Gly Thr Thr Thr Thr Thr Thr
465                        470                        475                        480

Gln Pro Thr Thr Thr Thr Thr Thr Ala Gly Asn Pro Gly Gly Thr Gly
      485                          490                          495

Val Ala Gln His Tyr Gly Gln Cys Gly Gly Ile Gly Trp Thr Gly Pro
      500                          505                          510

Thr Thr Cys Ala Ser Pro Tyr Thr Cys Gln Lys Leu Asn Asp Tyr Tyr
      515                        520                        525

Ser Gln Cys Leu
      530

```

<210> 7
<211> 25
<212> DNA
<213> *Aspergillus fumigatus*

<400> 7
gggcatgctg gcctccacct tctcc 25

<210> 8
<211> 30
<212> DNA
<213> *Aspergillus fumigatus*

<400> 8
gggtaatta actacaggca ctgagagtaa 30

<210> 9
<211> 1713
<212> DNA
<213> *Aspergillus fumigatus*

<400> 9
atgaagcacc ttgcattctc catgcattg actctactgt tgcctgccgt gcaggcccag 60
gagagcgtat ggggggaata tatcttctac ctctgactac cctaaagata tatgaagtag 120

```

cagaccgtat gggggcaatg catgtctctg ctgtcaactg aataagactg catcaactgc      120
tgatatgctt ctagggtggcg gccaaaggctg gtctggcccg acgagctgtg ttgccggcgc      180
agcctgtagc aactgaatc cctgtatgtt agatatcgtc ctgagtggag acttatactg      240
acttccttag actacgctca gtgtatcccg ggagccaccg cgacgtccac caccctcacg      300

acgacgacgg cggcgacgac gacatcccag accaccacca aacctaccac gactggtcca      360
actacatccg caccacccgt gaccgcatcc ggtaaccctt tcagcggcta ccagctgtat      420
gccaaacctt actactctc cgaggtccat actctggcca tgccttctct gccagctcg      480
ctgcagccca aggctagtgc tgttgctgaa gtgccctcat ttgtttggct gtaagtggcc      540
ttatcccaat actgagacca actctctgac agtcgtagcg acgttgccgc caagggtgcc      600
actatgggaa cctacctggc cgacattcag gccaaagaaca aggcgggcgc caacctcct      660
atcgtgtgta tcttcgtggt ctacgacttg ccggaccgtg actgcgccgc totggccagt      720
aatggcgagt actcaattgc caacaacggt gtggccaact acaaggcgta cattgacgcc      780
atcgtgtctc agctggtgaa gtactctgac gttcacacca tcctcgtcat cggtaggcgc      840
tacacctccg ttgcgcgccg cttttctctg acatcttgca gaacctgaca gcttggccaa      900
cctggtgacc aacctcaacg tcgccaaatg cgccaatgcg cagagcgccct acctggagtg      960
tgtcgactat gctctgaagc agctcaacct gcccaacgtc gccatgtacc tcgacgcagg     1020
tatgcctcac ttccgcgatt ctgtatccct tccagacact aactcatcag gccatgcggg     1080
ctggctcgga tggcccgcca acttggggcc cgccgcaaca ctcttcgcca aagtctacac     1140
cgacgcgggt tcccccgcgg ctgttcgtgg cctggccacc aacgtcgcca actacaacgc     1200
ctggtcgtct agtacctgcc cctcctacac ccaggagagc cccaactgcg acgagaagaa     1260
gtacatcaac gccatggcgc ctcttctcaa ggaagccggc ttcgatgcc acttcatcat     1320
ggatacctgt aagtgccttat tccaatcgcc gatgtgtgcc gactaatcaa tgtttcagcc     1380
cggaatggcg tccagcccac gaagcaaaac gcctggggtg actggtgcaa cgtcatcggc     1440
accggcttcg gtgttcgccc ctcgactaac accggcgatc cgctccagga tgcctttgtg     1500
tggatcaagc ccggtggaga gagtgatggc acgtccaact cgacttcccc ccggtatgac     1560
gcgcactgcg gatatagtga tgcctgcag cctgctcctg aggtggttac ttggttccag     1620
gtatgtcatc cattagccag atgagggata agtgactgac ggacctaggc ctactttgag     1680
cagcttctga ccaacgctaa ccgctccttt taa                                     1713

```

<210> 10

<211> 454

<212> PRT

<213> *Aspergillus fumigatus*

<400> 10

```

Met Lys His Leu Ala Ser Ser Ile Ala Leu Thr Leu Leu Leu Pro Ala
1           5           10           15

```

```

Val Gln Ala Gln Gln Thr Val Trp Gly Gln Cys Gly Gly Gln Gly Trp
20           25           30

```

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

[illegible]

<210> 11

<211> 37

<212> DNA

<213> *Aspergillus fumigatus*

<400> 11

actggattta ccatgaagca ccttgcattt tccatcg 37

<210> 12

$\langle 211 \rangle$ 34

<212> DNA

<213> *Aspergillus fumigatus*

<400> 12

tcacctctag ttaattaaaa ggacggggtta gcgt 34

<210> 13

<211> 799

<212> DNA

<213> Thermoascus aurantiacus

<400> 13

atgtcctttt ccaagataat tgctactgcc ggcgttcttg cctctgcttc tctagtggct 60

ggccatggct tcgttcagaa catcgtgatt gatggtaaaa agtatgtcat tgcaagacgc 120

acataagcgg caacagctga caatcgacag ttatggcggg tatctagtga accagtatcc 180

```

atacatgtcc aatcctccag aggtcatcgc ctggtctact acggcaactg atcttgatt      240
tgtggacggt actggatacc aaacccaga tatcatctgc cataggggcg ccaagcctgg      300
agccctgact gctccagtct ctccaggagg aactgttgag cttcaatgga ctccatggcc      360
tgattctcac catggcccag ttatcaacta ccttgctccg tgcaatggtg attgttccac      420
tgtggataag acccaattag aattcttcaa aattgccgag agcgggtctca tcaatgatga      480
caatcctcct gggatctggg cttcagacaa tctgatagca gccaacaaca gctggactgt      540
caccattcca accacaattg cacctggaaa ctatgttctg aggcattgaga ttattgctct      600
tcactcagct cagaaccagg atggtgcccc gaactatccc cagtgcacca atctgcaggt      660
cactggaggt ggttctgata accctgctgg aactcttgga acggcactct accacgatac      720
cgatcctgga attctgatca acatctatca gaaactttcc agctatatca tccctggtcc      780
tcctctgtat actggttaa                                          799

```

<210> 14

<211> 249

<212> PRT

<213> *Thermoascus aurantiacus*

<400> 14

```

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

Ser Leu Val Ala Gly His Gly Phe Val Gln Asn Ile Val Ile Asp Gly
          20           25           30

Lys Tyr Tyr Gly Gly Tyr Leu Val Asn Gln Tyr Pro Tyr Met Ser Asn
          35           40           45

Pro Pro Glu Val Ile Ala Trp Ser Thr Thr Ala Thr Asp Leu Gly Phe
          50           55           60

Val Asp Gly Thr Gly Tyr Gln Thr Pro Asp Ile Ile Cys His Arg Gly
65           70           75           80

Ala Lys Pro Gly Ala Leu Thr Ala Pro Val Ser Pro Gly Gly Thr Val
          85           90           95

Glu Leu Gln Trp Thr Pro Trp Pro Asp Ser His His Gly Pro Val Ile
          100          105          110

Asn Tyr Leu Ala Pro Cys Asn Gly Asp Cys Ser Thr Val Asp Lys Thr
          115          120          125

Gln Leu Glu Phe Phe Lys Ile Ala Glu Ser Gly Leu Ile Asn Asp Asp
          130          135          140

Asn Pro Pro Gly Ile Trp Ala Ser Asp Asn Leu Ile Ala Ala Asn Asn
145          150          155          160

Ser Trp Thr Val Thr Ile Pro Thr Thr Ile Ala Pro Gly Asn Tyr Val
          165          170          175

```

Leu Arg His Glu Ile Ile Ala Leu His Ser Ala Gln Asn Gln Asp Gly
 180 185 190

Ala Gln Asn Tyr Pro Gln Cys Ile Asn Leu Gln Val Thr Gly Gly Gly
 195 200 205

Ser Asp Asn Pro Ala Gly Thr Leu Gly Thr Ala Leu Tyr His Asp Thr
 210 215 220

Asp Pro Gly Ile Leu Ile Asn Ile Tyr Gln Lys Leu Ser Ser Tyr Ile
 225 230 235 240

Ile Pro Gly Pro Pro Leu Tyr Thr Gly
 245

<210> 15

<211> 3060

<212> DNA

<213> Aspergillus fumigatus

<400> 15

atgagattcg gttggctcga ggtggccgct ctgacggccg cttctgtagc caatgcccag	60
gtttgtgatg ctttcccgtc attgtttcgg atatagttga caatagtcac ggaaataatc	120
aggaattggc tttctctcca ccattctacc cttcgccttg ggctgatggc cagggagagt	180
gggcagatgc ccatcgacgc gccgtcgaga tcgtttctca gatgacactg gcggagaagg	240
ttaaccttac aacgggtact ggtggttg cgactttttt gttgacagtg agctttcttc	300
actgaccatc tacacagatg ggaaatggac cgatgcgtcg gtcaaaccgg cagcgttccc	360
aggtaagctt gcaattctgc aacaacgtgc aagtgtagtt gctaaaacgc ggtggtgcag	420
acttggatc aactggggtc tttgtggcca ggattccct ttgggtatcc gttctgtga	480
gctatacccg cggagtcttt cagtcttctg attatgtgct gatgattgtc tctgtatagc	540
tgacctcaac tccgccttcc ctgctggtac taatgtcgcc gcgacatggg acaagacact	600
cgcctacctt cgtggcaagg ccatgggtga ggaattcaac gacaagggcg tggacatttt	660
gctggggcct gctgctggtc ctctcgcaa ataccggac ggccggcagaa tctgggaagg	720
cttctctcct gatccggttc tcaactggtg acttttcgcc gaaactatca agggtatcca	780
agacgcgggt gtgattgcta ctgccaaaga ttacattctg aatgaacagg agcatttccg	840
acaggttggc gaggcccagg gatatggtta caacatcacg gagacgatca gctccaacgt	900
ggatgacaag accatgcacg agttgtacct ttggtgagta gttgacactg caaatgagga	960
ccttgattga ttgactgac ctggaatgca ggccctttgc agatgctgtg cgcggtgaaga	1020
ttttccgtag acttgacctc gcgacgaaga aatcgctgac gaaccatcgt agctggcggt	1080
ggcgtgtca tgtgttccca caatcaaata aacaacagct acggttgtca aaacagtcaa	1140
actctcaaca agctcctcaa ggtgagctg ggtctccaag gcttcgtcat gagtactgg	1200
agcgtcacc acagcgtgt cgccgtgcc ctgcgtgggt tggatatgtc gatgectgga	1260
gacatttctc tcgacgaagg actctccttc tggggcacga acctaaactgt cagtgttctt	1320
acacacacac ttccacacac caatctacac caatctacac ttactatact acacacacac	1380

```

aacgggacccg tcccgagccg ggggtgcgat gacatggctg tccgtaccat gaccggctac 1380
tacaagggtg gtctgtaccg tottcgtatt cccctaact tcagctcctg gaccgggat 1440
gagtaaggct gggagcattc tctgtctcc gagggagcct ggaccaagg gaacgacttc 1500
gtcaatgtgc agcgagtcga ctctcagatc atccgtgaga ttggtgccgc tagtacagt 1560
ctcttgaaga acacgggtgc tottcctttg accggcaagg aggttaaagt ggggtgtctc 1620
gggtgaagacg ctgggttccaa cccgtggggt gctaacggct gccccgaccg cggtgtgat 1680
aacggcactc ttgctatggc ctggggtagt ggtactgcc aattccctta cctgtcacc 1740
cccagcagg ctatccagg agaggtcatc agcaacggcg gcaatgtctt tctgtgact 1800
gataacgggg ctctcagcca gatggcagat gttgcatctc aatccagggt agtgccggct 1860
cttagaaaaa gaacgttctc tgaatgaagt tttttaacca ttgcgaacag cgtgtctttg 1920
gtgtttgtca acgcccactc tggagagggt ttcacagtg tcgacggcaa cgagggtgac 1980
cgcaaaaatc tcaactctgt gaagaacggc gaggcctga ttgacactgt tctcagccac 2040
tgcaacaaca cgattgtggt tattcacagt gttgggcccg tcttgatcga ccggtggtat 2100
gataacccca acgtcactgc catcatctgg gccggcttgc ccggtcagga gaggggcaac 2160
tccctggctg acgtgctcta tggccgctc aaccccagg ccaagacccc gttcacctgg 2220
ggcaagactc gggagtotta cggggctccc ttgctaccg agcctaaca tggcaatggt 2280

gctccccagg atgatttcaa cgagggcgctc ttcattgact accgtcactt tgacaagcgc 2340
aatgagaccc ccatttatga gtttggccat ggcttgagct acaccacctt tggttactct 2400
caccttcggg ttcaggccct caatagtctg agttcggcat atgtcccgac tagcggagag 2460
accaagcctg cgccaacctg tggtagatc ggtagtgcg ccgactacct gtatcccgag 2520
ggtctcaaaa gaattaccaa gtttatttac ccttggtcga actcgaccga cctcgaggat 2580
totctgacg acccgaacta cggtgggag gactcggagt acattcccga aggcgctagg 2640
gatgggtctc ctcaaccct cctgaaggct ggcggcgctc ctggtggtaa ccctaccctt 2700
tatcaggatc ttgttagggt gtccggcacc ataaccaaca ctggtaacgt cgccggttat 2760
gaagtcctc aattggtgag tgaccgcgat gttccttgcg ttgcaatttg gctaactcgc 2820
ttctagtatg tttcactggg cggaccgaac gagcctcggg tcgttctgcg caagtctgac 2880
cgaatcttcc tggctcctgg ggagcaaaag gtttgacca cgactotta cgtcgtgat 2940
ctcgccaatt gggatgtgga ggctcaggac tgggtcatca caaagtaccc caagaaagt 3000
cacgtcggca gctcctcgcg taagctgcct ctgagagcgc ctctgccccg tgtctactag 3060

```

<210> 16

<211> 863

<212> PRT

<213> *Aspergillus fumigatus*

<400> 16

```

Met Arg Phe Gly Trp Leu Glu Val Ala Ala Leu Thr Ala Ala Ser Val
1           5           10           15

```

```

Ala Asn Ala Gln Glu Leu Ala Phe Ser Pro Pro Phe Tyr Pro Ser Pro
20           25           30

```

Trp Ala Asp Gly Gln Gly Glu Trp Ala Asp Ala His Arg Arg Ala Val
 35 40 45
 Glu Ile Val Ser Gln Met Thr Leu Ala Glu Lys Val Asn Leu Thr Thr
 50 55 60
 Gly Thr Gly Trp Glu Met Asp Arg Cys Val Gly Gln Thr Gly Ser Val
 65 70 75 80
 Pro Arg Leu Gly Ile Asn Trp Gly Leu Cys Gly Gln Asp Ser Pro Leu
 85 90 95
 Gly Ile Arg Phe Ser Asp Leu Asn Ser Ala Phe Pro Ala Gly Thr Asn
 100 105 110
 Val Ala Ala Thr Trp Asp Lys Thr Leu Ala Tyr Leu Arg Gly Lys Ala
 115 120 125
 Met Gly Glu Glu Phe Asn Asp Lys Gly Val Asp Ile Leu Leu Gly Pro
 130 135 140
 Ala Ala Gly Pro Leu Gly Lys Tyr Pro Asp Gly Gly Arg Ile Trp Glu
 145 150 155 160
 Gly Phe Ser Pro Asp Pro Val Leu Thr Gly Val Leu Phe Ala Glu Thr
 165 170 175
 Ile Lys Gly Ile Gln Asp Ala Gly Val Ile Ala Thr Ala Lys His Tyr
 180 185 190
 Ile Leu Asn Glu Gln Glu His Phe Arg Gln Val Gly Glu Ala Gln Gly
 195 200 205
 Tyr Gly Tyr Asn Ile Thr Glu Thr Ile Ser Ser Asn Val Asp Asp Lys
 210 215 220
 Thr Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg Ala
 225 230 235 240
 Gly Val Gly Ala Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr
 245 250 255
 Gly Cys Gln Asn Ser Gln Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu
 260 265 270
 Gly Phe Gln Gly Phe Val Met Ser Asp Trp Ser Ala His His Ser Gly
 275 280 285
 Val Gly Ala Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Ile
 290 295 300
 Ser Phe Asp Asp Gly Leu Ser Phe Trp Gly Thr Asn Leu Thr Val Ser
 305 310 315 320
 Val Leu Asn Gly Thr Val Pro Ala Trp Arg Val Asp Asp Met Ala Val
 325 330 335

Arg Ile Met Thr Ala Tyr Tyr Lys Val Gly Arg Asp Arg Leu Arg Ile
340 345 350

Pro Pro Asn Phe Ser Ser Trp Thr Arg Asp Glu Tyr Gly Trp Glu His
355 360 365

Ser Ala Val Ser Glu Gly Ala Trp Thr Lys Val Asn Asp Phe Val Asn
370 375 380

Val Gln Arg Ser His Ser Gln Ile Ile Arg Glu Ile Gly Ala Ala Ser
385 390 395 400

Thr Val Leu Leu Lys Asn Thr Gly Ala Leu Pro Leu Thr Gly Lys Glu
405 410 415

Val Lys Val Gly Val Leu Gly Glu Asp Ala Gly Ser Asn Pro Trp Gly
420 425 430

Ala Asn Gly Cys Pro Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met
435 440 445

Ala Trp Gly Ser Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu
450 455 460

Gln Ala Ile Gln Arg Glu Val Ile Ser Asn Gly Gly Asn Val Phe Ala
465 470 475 480

Val Thr Asp Asn Gly Ala Leu Ser Gln Met Ala Asp Val Ala Ser Gln
485 490 495

Ser Ser Val Ser Leu Val Phe Val Asn Ala Asp Ser Gly Glu Gly Phe
500 505 510

Ile Ser Val Asp Gly Asn Glu Gly Asp Arg Lys Asn Leu Thr Leu Trp
515 520 525

Lys Asn Gly Glu Ala Val Ile Asp Thr Val Val Ser His Cys Asn Asn
530 535 540

Thr Ile Val Val Ile His Ser Val Gly Pro Val Leu Ile Asp Arg Trp
545 550 555 560

Tyr Asp Asn Pro Asn Val Thr Ala Ile Ile Trp Ala Gly Leu Pro Gly
565 570 575

Gln Glu Ser Gly Asn Ser Leu Val Asp Val Leu Tyr Gly Arg Val Asn
580 585 590

Pro Ser Ala Lys Thr Pro Phe Thr Trp Gly Lys Thr Arg Glu Ser Tyr
595 600 605

Gly Ala Pro Leu Leu Thr Glu Pro Asn Asn Gly Asn Gly Ala Pro Gln
610 615 620

Asp Asp Phe Asn Glu Gly Val Phe Ile Asp Tyr Arg His Phe Asp Lys
625 630 635 640

Arg Asn Glu Thr Pro Ile Tyr Glu Phe Gly His Gly Leu Ser Tyr Thr
645 650 655

Thr Phe Gly Tyr Ser His Leu Arg Val Gln Ala Leu Asn Ser Ser Ser
660 665 670

Ser Ala Tyr Val Pro Thr Ser Gly Glu Thr Lys Pro Ala Pro Thr Tyr
675 680 685

Gly Glu Ile Gly Ser Ala Ala Asp Tyr Leu Tyr Pro Glu Gly Leu Lys
690 695 700

Arg Ile Thr Lys Phe Ile Tyr Pro Trp Leu Asn Ser Thr Asp Leu Glu
705 710 715 720

Asp Ser Ser Asp Asp Pro Asn Tyr Gly Trp Glu Asp Ser Glu Tyr Ile
725 730 735

Pro Glu Gly Ala Arg Asp Gly Ser Pro Gln Pro Leu Leu Lys Ala Gly
740 745 750

Gly Ala Pro Gly Gly Asn Pro Thr Leu Tyr Gln Asp Leu Val Arg Val
755 760 765

Ser Ala Thr Ile Thr Asn Thr Gly Asn Val Ala Gly Tyr Glu Val Pro
770 775 780

Gln Leu Tyr Val Ser Leu Gly Gly Pro Asn Glu Pro Arg Val Val Leu
785 790 795 800

Arg Lys Phe Asp Arg Ile Phe Leu Ala Pro Gly Glu Gln Lys Val Trp
805 810 815

Thr Thr Thr Leu Asn Arg Arg Asp Leu Ala Asn Trp Asp Val Glu Ala
820 825 830

Gln Asp Trp Val Ile Thr Lys Tyr Pro Lys Lys Val His Val Gly Ser
835 840 845

Ser Ser Arg Lys Leu Pro Leu Arg Ala Pro Leu Pro Arg Val Tyr
850 855 860

<210> 17

<211> 1415

<212> DNA

<213> Aspergillus fumigatus

<400> 17

atggtccatc tatcttcatt ggcagcagcc ctggctgctc tgcctctgta tgtttaccca	60
ctcagcagag gaggaacagc ttgacattg ctatagtgtat tatggagctg gcctgaacac	120
agcagccaaa gccaaaggac taaagtactt tgggtccgcc acggacaatc cagagctcac	180
ggactctgcg tatgtcgcgc aactgagcaa caccgatgat tttgggtcaaa tcacaccg	240
-----	300

```

aaactccatg aaggtttgct tacgtctgcc tccctggagc attgcctcaa aagctaattg      300
gttgttttgt ttggatagtg ggatgccacc gagccttctc agaattcttt ttcgttcgca      360
aatggagacg ccgtgggtcaa tctggcgaac aagaatggcc agctgatgcg atgccatact      420
ctggctctggc acagtcagct accgaactgg ggtatgtaaa cgtcttgtct attctcaaat      480
actctctaac agttgacagt ctctagcggg tcatggacca atgcgaccct tttggcggcc      540
atgaagaatc atatcaccaa tgtggttact cactacaagg ggaagtgcta cgctgggat      600
gttgtcaatg aaggtttggt gctccatcta tctcaatag ttcttttgaa actgacaagc      660
ctgtcaatct agccctgaac gaggcaggta ctttccgtaa ctctgtcttc taccagatca      720
tcggcccagc atacattcct attgcgttcg ccacggctgc tgccgcagat cccgacgtga      780
aactctacta caacgactac aacattgaat actcaggcgc caaagcgact gctgcgcaga      840
atatcgtcaa gatgatcaag gcctacggcg cgaagatcga cggcgtcggc ctccaggcac      900
actttatcgt cggcagcact ccgagtcagt cggatctgac gaccgtcttg aagggctaca      960
ctgctctcgg cgttgagggt gcctataccg aacttgacat ccgcatgcag ctgccctcga     1020
ccgccgcaaa gctggcccag cagtccactg acttccaagg cgtggccgca gcatgcgtta     1080
gcaccactgg ctgcgtgggt gtcactatct gggactggac cgacaagtac tctgggtccc     1140
ccagcgtgtt ccaaggctac ggcgccccat tgccttggga tgagaactat gtgaagaagc     1200
cagcgtacga tggcctgatg gcgggtcttg gagcaagcgg ctccggcacc acaacgacca     1260
ctactactac ttctactacg acaggaggtg cggaccctac tggagtcgct cagaaatggg     1320
gacagtgtgg cgggtattggc tggaccgggc caacaacttg tgtcagtggg accacttgcc     1380
aaaagctgaa tgactggtac tcacagtgcc tgtaa                                  1415

```

<210> 18

<211> 397

<212> PRT

<213> *Aspergillus fumigatus*

<400> 18

```

Met Val His Leu Ser Ser Leu Ala Ala Ala Leu Ala Ala Leu Pro Leu
1           5           10           15

```

```

Val Tyr Gly Ala Gly Leu Asn Thr Ala Ala Lys Ala Lys Gly Leu Lys
20           25           30

```

```

Tyr Phe Gly Ser Ala Thr Asp Asn Pro Glu Leu Thr Asp Ser Ala Tyr
35           40           45

```

```

Val Ala Gln Leu Ser Asn Thr Asp Asp Phe Gly Gln Ile Thr Pro Gly
50           55           60

```

```

Asn Ser Met Lys Trp Asp Ala Thr Glu Pro Ser Gln Asn Ser Phe Ser
65           70           75           80

```

```

Phe Ala Asn Gly Asp Ala Val Val Asn Leu Ala Asn Lys Asn Gly Gln
85           90           95

```

```

Leu Met Arg Cys His Thr Leu Val Trp His Ser Gln Leu Pro Asn Trp

```

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

<210> 19

<211> 2388

<212> DNA

<213> *Talaromyces emersonii*

<400> 19

atgatgactc ccacggcgat tctcaccgca gtggcgggcg tcctgcccac cgcgacatgg	60
gcacaggata accaaaccta tgccaattac tctgcgcagt ctgagccgga cctgtttccc	120
cggaccgtcg cgaccatcga cctgtccttc cccgactgtg agaatggccc gctcagcacg	180
aacctggtgt gcaacaaatc ggccgatccc tgggcccagag ctgaggccct catctcgctc	240
tttaccctcg aagagctgat taacaacacc cagaacaccg ctcttggcgt gccccgtttg	300
ggtctgcccc agtatcaggt gtggaatgaa gctctgcacg gactggaccg cgccaatttc	360
tccattcggg gcgaatacag ctggggccacg tccttcccca tgcccatcct gtcgatggcg	420
tccttcaacc ggaccctcat caaccagatt gcctccatca ttgcaacgca agcccggtgc	480
ttcaacaacg ccggccgtta cggccttgac agctatgcgc ccaacatcaa tggcttcgcg	540
agtccctctt ggggcccgtg acaggagacg cctggtgagg atgcgttctt cttgagttcc	600
acctatgcgt acgagtacat cacaggcctg cagggcggtg tcgacccaga gcatgtcaag	660
atcgtcgcga cggcgaagca cttcgccggc tatgatctgg agaactgggg caacgtctct	720
cggctggggt tcaatgctat catcacgcag caggatctct ccgagtacta caccctcag	780
ttcctggcgt ctgctcgata cgccaagacg cgcagcatca tgtctccta caatgcagtg	840
aatggagtcc caagctgtgc caactccttc ttctccaga cgtttctccg agaaaacttt	900
gacttcgttg acgacgggta cgtctcgtcg gattgcgacg ccgtctacaa cgtcttcaac	960
ccacacgggt acgcccttaa ccagtcggga gccgctgcgg actcgtcctt agcaggatcc	1020
gatatcgact gtggtcagac cttgccgtgg cacctgaatg agtccttcgt agaaggatac	1080
gtctcccgcg gtgatatcga gaaatccctc acccgtctct actcaaacct ggtgcgtctc	1140
ggctactttg acggcaacaa cagcgagtac cgcaacctca actggaacga cgtcgtgact	1200
acggacgcct ggaacatctc gtacgaggcc gcggtggaag gtatcaccct gctcaagaac	1260
gacggaacgc tgccgctgtc caagaaggtc cgcagcattg cgtcatcgg tccttggggc	1320
aatgccacgg tgcagatgca gggtaactac tatggaacgc caccgtatct gatcagtcg	1380
ctggaagccg ccaaggccag tgggttcacg gtcaactatg cattcggtag caacatctcg	1440
accgattcta ccagtggtt gcggaagcc atcgcgcgcg cgaagaagtc ggacgtgatc	1500
atctacgccg gtggtattga caacacgatc gaggcagagg gacaggaccg cacggatctc	1560
aagtggccgg ggaaccagct ggatctgatc gacgagctca gccagggtgg caagcccttg	1620
gtcgtcctgc agatggggcg tggccagggt gattcgtcgt cactcaaggc caacaagaat	1680
gtcaacgctc tgggtgtggg tggctatccc ggacagtcgg gtggtgcggc cctgtttgac	1740
atccttacgg gcaagcgtgc gccggccggt cgtctggtga gcacgcagta cccggccgag	1800
tatgcgacgc agttcccggc caacgacatg aacctgcgtc cgaacggcag caaccggga	1860

cagacataca tctggtacac gggcacgccc gtgtatgagt tcggccacgg tctgttctac 1920
 acggagttcc aggagtgggc tgcggcgggc acgaacaaga cgtcgacttt cgacattctg 1980
 gaccttttct ccacctctca tccgggatac gagtacatcg agcaggttcc gttcatcaac 2040
 gtgactgtgg acgtgaagaa cgtcggccac acgccatcgc cgtacacggg tctgttggtc 2100
 gcgaacacga cagccggggc caagccgtac ccgaacaaat ggctcgtcgg gttcgactgg 2160
 ctgccgacga tccagccggg cgagactgcc aagttgacga tcccggtgcc gttgggcgcg 2220
 attgctggg cggacgagaa cggcaacaag gtggtcttcc cgggcaacta cgaattggca 2280
 ctgaacaatg agcgatcggt agtgggtgctg ttcacgctga cgggcgatgc ggcgactcta 2340
 gagaaatggc ctttgtggga gcaggcgggt ccgggggtgc tgcagcaa 2388

<210> 20

<211> 796

<212> PRT

<213> Talaromyces emersonii

<400> 20

Met Met Thr Pro Thr Ala Ile Leu Thr Ala Val Ala Ala Leu Leu Pro
1 5 10 15

Thr Ala Thr Trp Ala Gln Asp Asn Gln Thr Tyr Ala Asn Tyr Ser Ser
20 25 30

Gln Ser Gln Pro Asp Leu Phe Pro Arg Thr Val Ala Thr Ile Asp Leu
35 40 45

Ser Phe Pro Asp Cys Glu Asn Gly Pro Leu Ser Thr Asn Leu Val Cys
50 55 60

Asn Lys Ser Ala Asp Pro Trp Ala Arg Ala Glu Ala Leu Ile Ser Leu
65 70 75 80

Phe Thr Leu Glu Glu Leu Ile Asn Asn Thr Gln Asn Thr Ala Pro Gly
85 90 95

Val Pro Arg Leu Gly Leu Pro Gln Tyr Gln Val Trp Asn Glu Ala Leu
100 105 110

His Gly Leu Asp Arg Ala Asn Phe Ser His Ser Gly Glu Tyr Ser Trp
115 120 125

Ala Thr Ser Phe Pro Met Pro Ile Leu Ser Met Ala Ser Phe Asn Arg
130 135 140

Thr Leu Ile Asn Gln Ile Ala Ser Ile Ile Ala Thr Gln Ala Arg Ala
145 150 155 160

Phe Asn Asn Ala Gly Arg Tyr Gly Leu Asp Ser Tyr Ala Pro Asn Ile
165 170 175

Asn Gly Phe Arg Ser Pro Leu Trp Gly Arg Gly Gln Glu Thr Pro Gly
180 185 190

Glu Asp Ala Phe Phe Leu Ser Ser Thr Tyr Ala Tyr Glu Tyr Ile Thr
 195 200 205
 Gly Leu Gln Gly Gly Val Asp Pro Glu His Val Lys Ile Val Ala Thr
 210 215 220
 Ala Lys His Phe Ala Gly Tyr Asp Leu Glu Asn Trp Gly Asn Val Ser
 225 230 235 240
 Arg Leu Gly Phe Asn Ala Ile Ile Thr Gln Gln Asp Leu Ser Glu Tyr
 245 250 255
 Tyr Thr Pro Gln Phe Leu Ala Ser Ala Arg Tyr Ala Lys Thr Arg Ser
 260 265 270
 Ile Met Cys Ser Tyr Asn Ala Val Asn Gly Val Pro Ser Cys Ala Asn
 275 280 285
 Ser Phe Phe Leu Gln Thr Leu Leu Arg Glu Asn Phe Asp Phe Val Asp
 290 295 300
 Asp Gly Tyr Val Ser Ser Asp Cys Asp Ala Val Tyr Asn Val Phe Asn
 305 310 315 320
 Pro His Gly Tyr Ala Leu Asn Gln Ser Gly Ala Ala Ala Asp Ser Leu
 325 330 335
 Leu Ala Gly Thr Asp Ile Asp Cys Gly Gln Thr Leu Pro Trp His Leu
 340 345 350
 Asn Glu Ser Phe Val Glu Gly Tyr Val Ser Arg Gly Asp Ile Glu Lys
 355 360 365
 Ser Leu Thr Arg Leu Tyr Ser Asn Leu Val Arg Leu Gly Tyr Phe Asp
 370 375 380
 Gly Asn Asn Ser Glu Tyr Arg Asn Leu Asn Trp Asn Asp Val Val Thr
 385 390 395 400
 Thr Asp Ala Trp Asn Ile Ser Tyr Glu Ala Ala Val Glu Gly Ile Thr
 405 410 415
 Leu Leu Lys Asn Asp Gly Thr Leu Pro Leu Ser Lys Lys Val Arg Ser
 420 425 430
 Ile Ala Leu Ile Gly Pro Trp Ala Asn Ala Thr Val Gln Met Gln Gly
 435 440 445
 Asn Tyr Tyr Gly Thr Pro Pro Tyr Leu Ile Ser Pro Leu Glu Ala Ala
 450 455 460
 Lys Ala Ser Gly Phe Thr Val Asn Tyr Ala Phe Gly Thr Asn Ile Ser
 465 470 475 480
 Thr Asp Ser Thr Gln Trp Phe Ala Glu Ala Ile Ala Ala Ala Lys Lys
 485 490 495

Ser Asp Val Ile Ile Tyr Ala Gly Gly Ile Asp Asn Thr Ile Glu Ala
 500 505 510

Glu Gly Gln Asp Arg Thr Asp Leu Lys Trp Pro Gly Asn Gln Leu Asp
 515 520 525

Leu Ile Glu Gln Leu Ser Gln Val Gly Lys Pro Leu Val Val Leu Gln
 530 535 540

Met Gly Gly Gly Gln Val Asp Ser Ser Ser Leu Lys Ala Asn Lys Asn
 545 550 555 560

Val Asn Ala Leu Val Trp Gly Gly Tyr Pro Gly Gln Ser Gly Gly Ala
 565 570 575

Ala Leu Phe Asp Ile Leu Thr Gly Lys Arg Ala Pro Ala Gly Arg Leu
 580 585 590

Val Ser Thr Gln Tyr Pro Ala Glu Tyr Ala Thr Gln Phe Pro Ala Asn
 595 600 605

Asp Met Asn Leu Arg Pro Asn Gly Ser Asn Pro Gly Gln Thr Tyr Ile
 610 615 620

Trp Tyr Thr Gly Thr Pro Val Tyr Glu Phe Gly His Gly Leu Phe Tyr
 625 630 635 640

Thr Glu Phe Gln Glu Ser Ala Ala Ala Gly Thr Asn Lys Thr Ser Thr
 645 650 655

Phe Asp Ile Leu Asp Leu Phe Ser Thr Pro His Pro Gly Tyr Glu Tyr
 660 665 670

Ile Glu Gln Val Pro Phe Ile Asn Val Thr Val Asp Val Lys Asn Val
 675 680 685

Gly His Thr Pro Ser Pro Tyr Thr Gly Leu Leu Phe Ala Asn Thr Thr
 690 695 700

Ala Gly Pro Lys Pro Tyr Pro Asn Lys Trp Leu Val Gly Phe Asp Trp
 705 710 715 720

Leu Pro Thr Ile Gln Pro Gly Glu Thr Ala Lys Leu Thr Ile Pro Val
 725 730 735

Pro Leu Gly Ala Ile Ala Trp Ala Asp Glu Asn Gly Asn Lys Val Val
 740 745 750

Phe Pro Gly Asn Tyr Glu Leu Ala Leu Asn Asn Glu Arg Ser Val Val
 755 760 765

Val Ser Phe Thr Leu Thr Gly Asp Ala Ala Thr Leu Glu Lys Trp Pro
 770 775 780

Leu Trp Glu Gln Ala Val Pro Gly Val Leu Gln Gln
 785 790 795

<210> 21
<211> 19
<212> PRT
<213> Thielavia terrestris

<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> X=I,L,M, OR V

<220>
<221> misc_feature
<222> (3)..(6)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (8)..(8)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> X=I,L,M, OR V

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (13)..(13)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> MISC_FEATURE
<222> (14)..(14)
<223> X=E OR Q

<220>
<221> misc_feature
<222> (15)..(18)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> MISC_FEATURE

<222> (19)..(19)

<223> X=H,N, OR Q

<400> 21

Xaa	Pro	Xaa	Xaa	Xaa	Xaa	Gly	Xaa	Tyr	Xaa	Xaa	Arg	Xaa	Xaa	Xaa	Xaa
1			5						10					15	

Xaa Xaa Xaa

<210> 22

<211> 20

<212> PRT

<213> Thielavia terrestris

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> X=I,L,M, OR V

<220>

<221> misc_feature

<222> (3)..(7)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (9)..(9)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (11)..(11)

<223> X=I,L,M, OR V

<220>

<221> misc_feature

<222> (12)..(12)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (14)..(14)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (15)..(15)

<223> X=E OR Q

<220>

<221> misc_feature

<222> (16)..(19)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (20)..(20)

<223> X=H,N, OR Q

<400> 22

Xaa	Pro	Xaa	Xaa	Xaa	Xaa	Xaa	Gly	Xaa	Tyr	Xaa	Xaa	Arg	Xaa	Xaa	Xaa
1				5					10					15	

Xaa	Xaa	Xaa	Xaa
			20

<210> 23

<211> 9

<212> PRT

<213> Thielavia terrestris

<220>

<221> misc_feature

<222> (2)..(2)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (5)..(7)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (8)..(8)

<223> X= Y OR W

<220>

<221> MISC_FEATURE

<222> (9)..(9)

<223> X= A,I,L,M OR V

<220>

<221> MISC_FEATURE

<222> (9)..(9)

<223> X= A,I,L,M OR V

<400> 23

His	Xaa	Gly	Pro	Xaa	Xaa	Xaa	Xaa	Xaa
-----	-----	-----	-----	-----	-----	-----	-----	-----

```

His Xaa Gly P10 Xaa Xaa Xaa Xaa Xaa
1                      5

```

<210> 24

<211> 10

<212> PRT

<213> Thielavia terrestris

<220>

<221> misc_feature

<222> (2)..(3)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (6)..(8)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (9)..(9)

<223> X= Y OR W

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> X= A,I,L,M OR V

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> X= A,I,L,M OR V

<400> 24

```

His Xaa Xaa Gly Pro Xaa Xaa Xaa Xaa Xaa
1                      5                      10

```

<210> 25

<211> 11

<212> PRT

<213> Thielavia terrestris

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> X= E OR Q

<220>

<221> misc_feature

<222> (2)..(2)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (4)..(5)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (7)..(7)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (8)..(8)

<223> X= E,H,Q OR N

<220>

<221> MISC_FEATURE

<222> (9)..(9)

<223> X=F,I,L, OR V

<220>

<221> misc_feature

<222> (10)..(10)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (11)..(11)

<223> X=I,L,OR V

<220>

<221> MISC_FEATURE

<222> (11)..(11)

<223> X=I,L,OR V

<400> 25

Xaa	Xaa	Tyr	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa
1				5						10

<210> 26

<211> 9

<212> PRT

<213> Thielavia terrestris

<220>

<221> misc_feature
 <222> (2)..(2)
 <223> Xaa can be any naturally occurring amino acid

<220>
 <221> misc_feature
 <222> (5)..(7)
 <223> Xaa can be any naturally occurring amino acid

<220>
 <221> MISC_FEATURE
 <222> (8)..(8)
 <223> X= Y OR W

<220>
 <221> MISC_FEATURE
 <222> (9)..(9)
 <223> X= A,I,L,M OR V

<400> 26
 His Xaa Gly Pro Xaa Xaa Xaa Xaa Xaa
 1 5

<210> 27
 <211> 10
 <212> PRT
 <213> Thielavia terrestris

<220>
 <221> misc_feature
 <222> (2)..(3)
 <223> Xaa can be any naturally occurring amino acid

<220>
 <221> misc_feature
 <222> (6)..(8)
 <223> Xaa can be any naturally occurring amino acid

<220>
 <221> MISC_FEATURE
 <222> (9)..(9)
 <223> X= Y OR W

<220>
 <221> MISC_FEATURE
 <222> (10)..(10)
 <223> X= A,I,L,M OR V

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> X= A,I,L,M OR V

<400> 27

His	Xaa	Xaa	Gly	Pro	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10

<210> 28

<211> 11

<212> PRT

<213> Thielavia terrestris

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> X= E OR Q

<220>

<221> misc_feature

<222> (2)..(2)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (4)..(5)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (7)..(7)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (8)..(8)

<223> X= E,H,Q OR N

<220>

<221> MISC_FEATURE

<222> (9)..(9)

<223> X=F,I,L, OR V

<220>

<221> misc_feature

<222> (10)..(10)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (11)..(11)

<223> X=I,L,OR V

<400> 28

Xaa	Xaa	Tyr	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10	

<210> 29

<211> 19

<212> PRT

<213> Thielavia terrestris

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> X=I,L,M OR V

<220>

<221> misc_feature

<222> (3)..(6)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (8)..(8)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> X=I,L,M OR V

<220>

<221> misc_feature

<222> (11)..(11)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (13)..(13)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (14)..(14)

<223> X= E OR Q

<220>

<221> misc_feature

<222> (15)..(17)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (19)..(19)

<223> X= H,N, OR Q

<400> 29

Xaa	Pro	Xaa	Xaa	Xaa	Xaa	Gly	Xaa	Tyr	Xaa	Xaa	Arg	Xaa	Xaa	Xaa	Xaa
1			5						10					15	

Xaa Ala Xaa

<210> 30

<211> 20

<212> PRT

<213> Thielavia terrestris

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> X=I,L,M OR V

<220>

<221> misc_feature

<222> (3)..(7)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (9)..(9)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (11)..(11)

<223> X=I,L,M OR V

<220>

<221> misc_feature

<222> (12)..(12)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (14)..(14)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (15)..(15)

<223> X=E OR Q

<220>

<221> misc_feature

<222> (16)..(18)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> MISC_FEATURE

<222> (20)..(20)

<223> X= H,N, OR Q

<400> 30

Xaa	Pro	Xaa	Xaa	Xaa	Xaa	Gly	Xaa	Tyr	Xaa	Xaa	Arg	Xaa	Xaa	Xaa
1			5					10				15		

Xaa	Xaa	Ala	Xaa
		20	

<210> 31

<211> 2394

<212> DNA

<213> Trichoderma reesei

<400> 31

atggtgaata	acgcagctct	tctcgccgcc	ctgtcggtct	tcctgccac	ggcctggcg	60
cagaacaatc	aaacatacgc	caactactct	gctcagggcc	agcctgatct	ctaccccgag	120
acacttgcca	cgctcacact	ctcgttcccc	gactgcgaac	atggccccct	caagaacaat	180
ctcgtctgtg	actcatcggc	cggtatgta	gagcgagccc	aggccctcat	ctcgtctctc	240
accctcgagg	agctcattct	caacacgcaa	aactcgggcc	ccggcggtgc	tgcctgggt	300
cttcggaact	accaagtctg	gaatgaggct	ctgcacggct	tggaccgcgc	caacttcgcc	360
accaagggcg	gccagttcga	atgggcgacc	tcgttcccca	tgcccatcct	cactacggcg	420
gccctcaacc	gcacattgat	ccaccagatt	gccgacatca	tctcgacca	agctcgagca	480
ttcagcaaca	gcggcggtta	cggctctcgac	gtctatgcgc	caaacgtcaa	tggtttccga	540
agccccctct	ggggccgtgg	ccaggagacg	cccggcgaag	acgccttttt	cctcagctcc	600
gcctatactt	acgagtacat	cacgggcatc	cagggtggcg	tcgaccctga	gcacctcaag	660
gttgccgcc	cggtgaagca	ctttgccgga	tacgacctcg	agaactggaa	caaccagtc	720
cctctcggtt	tcgagcccat	cataactcag	caggacctct	cgaatacta	cactcccgag	780

```

cgtctcgggc cggagccat caccctccag caggaccccc cggatcacta cactccccag      700
ttctctcgctg cggcccggtta tgcaaagtca cgcagcttga tgtgcgcata caactccgtc      840
aacggcgtgc ccagctgtgc caacagcttc ttcttgcala cgcttttgcg cgagagctgg      900
ggcttccccg aatggggata cgtctcgtcc gattgcgatg ccgtctacaa cgttttcaac      960
cctcatgact acgccagcaa ccagtcgtca gccgccgcca gctcaactgcg agccggcacc     1020
gatatcgact gcggtcagac ttaccctggg cacctcaacg agtcctttgt ggccggcgaa     1080
gtctcccgcg gcgagatcga gcggtccgtc acccgtctgt acgccaacct cgtccgtctc     1140
ggatacttcg acaagaagaa ccagtaccgc tcgctcgggtt ggaaggatgt cgtcaagact     1200
gatgcctgga acatctcgta cgaggctgct gttgagggca tcgtcctgct caagaacgat     1260
ggcactctcc ctctgtccaa gaaggcgccg agcattgctc tgatcggacc atggggccaa     1320
gccacaaccc aaatgcaagg caactactat ggccctgccc catacctcat cagccctctg     1380
gaagctgcta agaaggccgg ctatcacgtc aactttgaac tcggcacaga gatcgccggc     1440
aacagcacca ctggctttgc caaggccatt gctgccgcca agaagtcgga tgccatcatc     1500
tacctcggtg gaattgacaa caccattgaa caggagggcg ctgaccgcac ggacattgct     1560

tggcccggtg atcagctgga tctcatcaag cagctcagcg aggtcggcaa accccttgct     1620
gtcctgcaaa tgggcggtgg tcaggtagac tcctcctcgc tcaagagcaa caagaagtc     1680
aactccctcg tctggggcgg atatcccggc cagtcgggag gcgttgccct ctctgacatt     1740
ctctctggca agcgtgctcc tgccggccga ctggtcacca ctcagtaccc ggctgagtat     1800
gttcaccaat tccccagaa tgacatgaac ctccgaccog atggaaagtc aaaccctgga     1860
cagacttaca tctggtacac cggcaaaccc gtctacagat ttggcagtgg tctctttctac     1920
accaccttca aggagactct cgccagccac cccaagagcc tcaagttcaa cacctcatcg     1980
atcctctctg ctctcaccg cggatacact tacagcgagc agattcccgct cttcaccttc     2040
gaggccaaca tcaagaactc gggcaagacg gagtcccat atacggccat gctgtttgtt     2100
cgcacaagca acgctggccc agccccgtac ccgaacaagt ggctcgtcgg attcgaccga     2160
cttgccgaca tcaagcctgg tcaactcttc aagctcagca tccccatccc tgtcagtgtc     2220
ctcgcccgctg ttgattctca cggaaaccgg attgtatacc ccggcaagta tgagctagcc     2280
ttgaacaccg acgagtctgt gaagcttgag tttgagttgg tgggagaaga ggtaacgatt     2340
gagaactggc cgttgaggga gcaacagatc aaggatgcta cacctgacgc ataa         2394

```

<210> 32

<211> 797

<212> PRT

<213> Trichoderma reesei

<400> 32

```

Met Val Asn Asn Ala Ala Leu Leu Ala Ala Leu Ser Ala Leu Leu Pro
1           5           10           15

```

```

Thr Ala Leu Ala Gln Asn Asn Gln Thr Tyr Ala Asn Tyr Ser Ala Gln
20           25           30

```

Gly Gln Pro Asp Leu Tyr Pro Glu Thr Leu Ala Thr Leu Thr Leu Ser
 35 40 45
 Phe Pro Asp Cys Glu His Gly Pro Leu Lys Asn Asn Leu Val Cys Asp
 50 55 60
 Ser Ser Ala Gly Tyr Val Glu Arg Ala Gln Ala Leu Ile Ser Leu Phe
 65 70 75 80
 Thr Leu Glu Glu Leu Ile Leu Asn Thr Gln Asn Ser Gly Pro Gly Val
 85 90 95
 Pro Arg Leu Gly Leu Pro Asn Tyr Gln Val Trp Asn Glu Ala Leu His
 100 105 110
 Gly Leu Asp Arg Ala Asn Phe Ala Thr Lys Gly Gly Gln Phe Glu Trp
 115 120 125
 Ala Thr Ser Phe Pro Met Pro Ile Leu Thr Thr Ala Ala Leu Asn Arg
 130 135 140
 Thr Leu Ile His Gln Ile Ala Asp Ile Ile Ser Thr Gln Ala Arg Ala
 145 150 155 160
 Phe Ser Asn Ser Gly Arg Tyr Gly Leu Asp Val Tyr Ala Pro Asn Val
 165 170 175
 Asn Gly Phe Arg Ser Pro Leu Trp Gly Arg Gly Gln Glu Thr Pro Gly
 180 185 190
 Glu Asp Ala Phe Phe Leu Ser Ser Ala Tyr Thr Tyr Glu Tyr Ile Thr
 195 200 205
 Gly Ile Gln Gly Gly Val Asp Pro Glu His Leu Lys Val Ala Ala Thr
 210 215 220
 Val Lys His Phe Ala Gly Tyr Asp Leu Glu Asn Trp Asn Asn Gln Ser
 225 230 235 240
 Arg Leu Gly Phe Asp Ala Ile Ile Thr Gln Gln Asp Leu Ser Glu Tyr
 245 250 255
 Tyr Thr Pro Gln Phe Leu Ala Ala Ala Arg Tyr Ala Lys Ser Arg Ser
 260 265 270
 Leu Met Cys Ala Tyr Asn Ser Val Asn Gly Val Pro Ser Cys Ala Asn
 275 280 285
 Ser Phe Phe Leu Gln Thr Leu Leu Arg Glu Ser Trp Gly Phe Pro Glu
 290 295 300
 Trp Gly Tyr Val Ser Ser Asp Cys Asp Ala Val Tyr Asn Val Phe Asn
 305 310 315 320
 Pro His Asp Tyr Ala Ser Asn Gln Ser Ser Ala Ala Ala Ser Ser Leu
 325 330 335

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

Ala Tyr Thr Gly Lys Pro Val Tyr Glu Phe Gly Ser Gly Leu Phe Tyr
625 630 635 640

Thr Thr Phe Lys Glu Thr Leu Ala Ser His Pro Lys Ser Leu Lys Phe
645 650 655

Asn Thr Ser Ser Ile Leu Ser Ala Pro His Pro Gly Tyr Thr Tyr Ser
660 665 670

Glu Gln Ile Pro Val Phe Thr Phe Glu Ala Asn Ile Lys Asn Ser Gly
675 680 685

Lys Thr Glu Ser Pro Tyr Thr Ala Met Leu Phe Val Arg Thr Ser Asn
690 695 700

Ala Gly Pro Ala Pro Tyr Pro Asn Lys Trp Leu Val Gly Phe Asp Arg
705 710 715 720

Leu Ala Asp Ile Lys Pro Gly His Ser Ser Lys Leu Ser Ile Pro Ile
725 730 735

Pro Val Ser Ala Leu Ala Arg Val Asp Ser His Gly Asn Arg Ile Val
740 745 750

Tyr Pro Gly Lys Tyr Glu Leu Ala Leu Asn Thr Asp Glu Ser Val Lys
755 760 765

Leu Glu Phe Glu Leu Val Gly Glu Glu Val Thr Ile Glu Asn Trp Pro
770 775 780

Leu Glu Glu Gln Gln Ile Lys Asp Ala Thr Pro Asp Ala
785 790 795

<210> 33

<211> 1849

<212> DNA

<213> Trichoderma reesei

<400> 33

tgccatttct gacctggata ggttttccta tggtcattcc tataagagac acgctctttc	60
gtcggcccggt agatatcaga ttggtattca gtcgcacaga cgaagggtgag ttgatcctcc	120
aacatgagtt ctatgagccc ccccttgcc ccccccggt caccttgacc tgcaatgaga	180
atcccacett ttacaagagc atcaagaagt attaatggcg ctgaatagcc tetgctcgat	240
aatatctccc cgtcatcgac aatgaacaag tccgtggctc cattgctgct tgcagcgtcc	300
atactatatg gcggcgccgt cgcacagcag actgtctggg gccagtgtgg aggtattggt	360
tggagcggac ctacgaattg tgctcctggc tcagcttggt cgacctcaa tccttattat	420
gcgcaatgta ttccgggagc cactactatc accacttoga cccggccacc atccggtcca	480
accaccacca ccagggtac ctcaacaagc tcatcaactc caccacgag ctctggggtc	540
cgatttgccg gcgttaacat cgcgggtttt gactttggct gtaccacaga gtgagtaccc	600
ttgtttcctg gtgttgctgg ctggttgggc gggatatacag cgaagcggac gcaagaacac	660

```

cgccggtccg ccaccatcaa gatgtgggtg gtaagcggcg gtgttttgta caactacctg      720
acagctcact caggaatga gaattaatgg aagtcttggt acagtggcac ttgcgttacc      780
togaaggttt atcctccggt gaagaacttc accggctcaa acaactaccc cgatggcatc      840
ggccagatgc agcacttcgt caacgaggac gggatgacta ttttccgctt acctgtcgga      900
tggcagtacc tcgtcaacaa caatttgggc ggcaatcttg attccacgag catttccaag      960
tatgatcagc ttgttcaggg gtgcctgtct ctgggcgcat actgcatcgt cgacatccac     1020
aattatgctc gatggaacgg tgggatcatt ggtcaggggc gccctactaa tgctcaattc     1080
acgagccttt ggtcgcagtt ggcacaaaag tacgcctctc agtcgagggg gtgggttcggc     1140
atcatgaatg agccccacga cgtgaacatc aacacctggg ctgccacggt ccaagagggt     1200
gtaaccgcaa tccgcaacgc tgggtgctacg tcgcaattca tctctttgcc tggaaatgat     1260
tggcaatctg ctggggcttt catatccgat ggcagtgcag ccgccctgtc tcaagtcacg     1320
aaccgggatg ggtcaacaac gaatctgatt tttgacgtgc acaaatactt ggactcagac     1380
aactccggta ctacgcgga atgtactaca aataacattg acggcgctt ttctccgctt     1440
gccacttggc tccgacagaa caatcgccag gctatcctga cagaaaccgg tgggtggcaac     1500
gttcagtctt gcatacaaga catgtgccag caaatccaat atctcaacca gaactcagat     1560
gtctatcttg gctatgttgg ttggggtgcc ggatcatttg atagcacgta tgtcctgacg     1620
gaaacaccga ctggcagtgg taactcatgg acggacacat ccttggtcag ctcgtgtctc     1680
gcaagaaagt agcactctga gctgaatgca gaagcctcgc caacgtttgt atctcgctat     1740
caaacatagt agctactcta tgaggctgtc tgttctcgat ttcagcttta tatagtttca     1800
tcaaacagta catattccct ctgtggccac gcaaaaaaaaa aaaaaaaaaa     1849

```

<210> 34

<211> 418

<212> PRT

<213> Trichoderma reesei

<400> 34

```

Met Asn Lys Ser Val Ala Pro Leu Leu Ala Ala Ser Ile Leu Tyr
1           5           10           15

Gly Gly Ala Val Ala Gln Gln Thr Val Trp Gly Gln Cys Gly Gly Ile
          20           25           30

Gly Trp Ser Gly Pro Thr Asn Cys Ala Pro Gly Ser Ala Cys Ser Thr
          35           40           45

Leu Asn Pro Tyr Tyr Ala Gln Cys Ile Pro Gly Ala Thr Thr Ile Thr
          50           55           60

Thr Ser Thr Arg Pro Pro Ser Gly Pro Thr Thr Thr Arg Ala Thr
65           70           75           80

Ser Thr Ser Ser Ser Thr Pro Pro Thr Ser Ser Gly Val Arg Phe Ala
          85           90           95

```

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

Ala Gly Ser Phe Asp Ser Thr Tyr Val Leu Thr Glu Thr Pro Thr Gly
 385 390 395 400

Ser Gly Asn Ser Trp Thr Asp Thr Ser Leu Val Ser Ser Cys Leu Ala
 405 410 415

Arg Lys

REFERENCES CITED IN THE DESCRIPTION

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

Patent documents cited in the description

- [US5861271A \[0004\]](#)
- [WO02095014A \[0051\]](#)
- [WO2002095014A \[0051\] \[0252\]](#)
- [WO9517413A \[0079\]](#)
- [WO9522625A \[0079\]](#)
- [US5223409A \[0079\]](#)
- [WO9206204A \[0079\]](#)
- [WO9943835A \[0108\]](#)
- [WO9600787A \[0109\] \[0157\]](#)
- [WO0056900A \[0109\] \[0109\] \[0109\]](#)
- [US6011147A \[0109\]](#)
- [WO9425612A \[0116\]](#)
- [WO9533836A \[0127\]](#)
- [WO2010039889A \[0134\]](#)
- [WO0024883A \[0140\] \[0140\]](#)
- [EP238023A \[0157\]](#)
- [WO9114772A \[0173\]](#)
- [US6395966B \[0177\]](#)
- [US7151204B \[0177\] \[0179\]](#)
- [WO9015861A \[0190\]](#)
- [WO2010096673A \[0190\]](#)
- [US20020164730A \[0207\] \[0216\]](#)
- [WO2006110891A \[0211\]](#)

- [WO2006110899A \[0211\]](#)
- [WO2006110900A \[0211\]](#)
- [WO2006110901A \[0211\]](#)
- [WO2006032282A \[0213\]](#)
- [WO9117243A \[0247\]](#)
- [WO9117244A \[0247\]](#)
- [WO9105039A \[0249\]](#)
- [WO9315186A \[0249\]](#)
- [US5275944A \[0249\]](#)
- [WO9602551A \[0249\]](#)
- [US5536655A \[0249\]](#)
- [WO0070031A \[0249\]](#)
- [WO05093050A \[0249\] \[0249\] \[0249\]](#)
- [WO2011059740A \[0251\]](#)
- [WO2009042871A \[0251\]](#)
- [WO2010141325A \[0251\]](#)
- [WO2006074435A \[0251\]](#)
- [WO2010057086A \[0251\]](#)
- [WO2005047499A \[0252\] \[0390\] \[0412\]](#)
- [WO2007019442A \[0252\] \[0252\]](#)
- [WO2010088387A \[0252\]](#)
- [WO2011035029A \[0252\]](#)
- [WO2008057637A \[0253\] \[0253\]](#)
- [WO9813465A \[0255\]](#)
- [WO98015619A \[0255\]](#)
- [WO98015633A \[0255\]](#)
- [WO9906574A \[0255\]](#)
- [WO9910481A \[0255\]](#)
- [WO99025847A \[0255\]](#)
- [WO99031255A \[0255\]](#)
- [WO2002101078A \[0255\]](#)
- [WO2003027306A \[0255\]](#)
- [WO2003052054A \[0255\]](#)
- [WO2003052055A \[0255\]](#)
- [WO2003052056A \[0255\]](#)
- [WO2003052057A \[0255\]](#)
- [WO2003052118A \[0255\]](#)
- [WO2004016760A \[0255\]](#)
- [WO2004043980A \[0255\]](#)
- [WO2004048592A \[0255\]](#)
- [WO2005001065A \[0255\]](#)
- [WO2005028636A \[0255\]](#)
- [WO2005093050A \[0255\]](#)
- [WO2005093073A \[0255\]](#)

- [WO2006074005A](#) [0255]
- [WO2006117432A](#) [0255]
- [WO2007071818A](#) [0255]
- [WO2007071820A](#) [0255]
- [WO2008008070A](#) [0255]
- [WO2008008793A](#) [0255]
- [US5457046A](#) [0255]
- [US5648263A](#) [0255]
- [US5686593A](#) [0255]
- [WO2005074647A](#) [0261]
- [WO2008148131A](#) [0261]
- [WO2011035027A](#) [0261]
- [WO2005074656A](#) [0261] [0411]
- [WO2010065830A](#) [0261]
- [WO2007089290A](#) [0261]
- [WO2009085935A](#) [0261]
- [WO2009085859A](#) [0261]
- [WO2009085864A](#) [0261]
- [WO2009085868A](#) [0261]
- [WO2010138754A](#) [0261]
- [WO2011005867A](#) [0261]
- [WO2011039319A](#) [0261]
- [WO2011041397A](#) [0261]
- [WO2011041504A](#) [0261]
- [WO2008151043A](#) [0262]
- [WO9421785A](#) [0274]
- [WO2006078256A](#) [0274] [0414]
- [WO2011041405A](#) [0274]
- [WO2010126772A](#) [0274]
- [WO2009079210A](#) [0274]
- [WO2011057083A](#) [0274]
- [WO2010108918A](#) [0276]
- [WO2009073709A](#) [0276]
- [WO2005001036A](#) [0276]
- [WO2010014880A](#) [0276]
- [WO2009042846A](#) [0276]
- [WO2009076122A](#) [0277]
- [WO2009127729A](#) [0277]
- [WO2010053838A](#) [0277]
- [WO2010065448A](#) [0277]
- [WO2006114094A](#) [0278] [0278]
- [WO2009073383A](#) [0278]
- [WO2010014706A](#) [0279]
- [WO2009068565A](#) [0279]

- [WO2003062430A \[0296\]](#)
- [US5681862A \[0339\]](#)
- [US5575993A \[0339\]](#)
- [US5256252A \[0339\]](#)
- [US5827719A \[0348\]](#)
- [WO9009446A \[0348\]](#)
- [WO9414964A \[0348\]](#)
- [WO9403578A \[0348\] \[0348\]](#)
- [WO0034450A \[0348\]](#)
- [WO0192502A \[0348\]](#)
- [WO8803947A \[0349\]](#)
- [WO9100345A \[0349\]](#)
- [EP415296A \[0349\]](#)
- [WO0053843A \[0352\]](#)
- [US6066486A \[0352\]](#)
- [JP2000080581B \[0352\]](#)
- [EP730641A \[0352\]](#)
- [WO0198469A \[0352\]](#)
- [EP719337A \[0352\]](#)
- [EP765394A \[0352\]](#)
- [EP767836A \[0352\]](#)
- [EP763115A \[0352\]](#)
- [EP788547A \[0352\]](#)
- [EP705327A \[0352\]](#)
- [WO9856899A \[0352\]](#)
- [EP677102A \[0352\]](#)
- [EP781328A \[0352\]](#)
- [EP707637A \[0352\]](#)
- [WO2003035972A \[0353\]](#)
- [WO200240694A \[0356\] \[0375\]](#)
- [WO2005042735A \[0371\]](#)
- [PE03860005710 \[0374\] \[0375\]](#)
- [EP0238023A \[0375\]](#)
- [WO2004099228A \[0392\]](#)
- [WO9961651A \[0393\]](#)
- [WO2000039322A \[0414\]](#)
- [WO2003070956A \[0416\]](#)
- [WO2011057140A \[0424\] \[0425\]](#)
- [US61480245B \[0431\]](#)

Non-patent literature cited in the description

- **FÄHNRICHIRRGANG**Biotechnology Letters, 1984, vol. 6, 4251-256 [0004]
- **DE VRIES**J. Bacteriol., 1998, vol. 180, 243-249 [0014]
- **VENTURI et al.**Extracellular beta-D-glucosidase from Chaetomium thermophilum var. coprophilum: production, purification and some biochemical propertiesJ. Basic Microbiol., 2002, vol. 42, 55-66 [0015]
- **TEERI**Crystalline cellulose degradation: New insight into the function of cellobiohydrolasesTrends in Biotechnology, 1997, vol. 15, 160-167 [0019]
- **TEERI et al.**Trichoderma reesei cellobiohydrolases: why so efficient on crystalline cellulose?Biochem. Soc. Trans., 1998, vol. 26, 173-178 [0019]
- **LEVER et al.**Anal. Biochem., 1972, vol. 47, 273-279 [0019]
- **VAN TILBEURGH et al.**FEBS Letters, 1982, vol. 149, 152-156 [0019]
- **VAN TILBEURGHCLAEYSSSENS**FEBS Letters, 1985, vol. 187, 283-288 [0019]
- **TOMME et al.**Eur. J. Biochem., 1988, vol. 170, 575-581 [0019]
- **ZHANG et al.**Outlook for cellulase improvement: Screening and selection strategiesBiotechnology Advances, 2006, vol. 24, 452-481 [0021]
- **GHOSE**Measurement of cellulase activitiesPure Appl. Chem., 1987, vol. 59, 257-68 [0021]
- **WISELOGEL et al.**Handbook on BioethanolTaylor & Francis19950000105-118 [0024]
- **WYMAN**Bioresource Technology, 1994, vol. 50, 3-16 [0024]
- **LYND**Applied Biochemistry and Biotechnology, 1990, vol. 24/25, 695-719 [0024]
- Recent Progress in Bioconversion of Lignocellulosics**MOSIER et al.**Advances in Biochemical Engineering/BiotechnologySpringer-Verlag19990000vol. 65, 23-40 [0024]
- **ZHANG et al.**Biotechnology Advances, 2006, vol. 24, 452-481 [0033]
- **GHOSE**Pure and Appl. Chem., 1987, vol. 59, 257-268 [0033]
- **HENRISSAT B.**A classification of glycosyl hydrolases based on amino-acid sequence similaritiesBiochem. J., 1991, vol. 280, 309-316 [0037] [0254]
- **HENRISSAT B.BAIROCH A.**Updating the sequence-based classification of glycosyl hydrolasesBiochem. J., 1996, vol. 316, 695-696 [0037] [0254]
- **SHALLOM, D.SHOHAM, Y.**Microbial hemicellulasesCurrent Opinion In Microbiology, 2003, vol. 6, 3219-228 [0040]
- **GHOSEBISARIA**Pure & Appl. Chem., 1987, vol. 59, 1739-1752 [0040]
- **NIELSEN et al.**Protein Engineering, 1997, vol. 10, 1-6 [0045] [0380]
- **NEEDLEMANWUNSCH**J. Mol. Biol., 1970, vol. 48, 443-453 [0055] [0381]
- **RICE et al.**EMBOSS: The European Molecular Biology Open Software SuiteTrends Genet., 2000, vol. 16, 276-277 [0055]
- **RICE et al.**EMBOSS: The European Molecular Biology Open Software Suite, 2000, [0056]
- **EBRINGEROVA et al.**Adv. Polym. Sci., 2005, vol. 186, 1-67 [0061]
- **BIELYPUCHARD**Recent progress in the assays of xylanolytic enzymesJournal of the Science of Food and Agriculture, 2006, vol. 86, 111636-1647 [0063]
- **SPANIKOVABIELY**Glucuronoyl esterase - Novel carbohydrate esterase produced by Schizophyllum communeFEBS Letters, 2006, vol. 580, 194597-4601 [0063]
- **HERRMANNVRSANSKAJURICKOVAHIRSCHBIELYKUBICEK**The beta-D-xylosidase of

Trichoderma reesei is a multifunctional beta-D-xylan xylohydrolase *Biochemical Journal*, 1997, vol. 321, 375-381 [0063]

- **BAILEYBIELYPOUTANEN** Interlaboratory testing of methods for assay of xylanase activity *Journal of Biotechnology*, 1992, vol. 23, 3257-270 [0064]
- **LEVERA** A new reaction for colorimetric determination of carbohydrates *Anal. Biochem.*, 1972, vol. 47, 273-279 [0065]
- **SAMBROOK et al.** *Molecular Cloning, A Laboratory Manual* Cold Spring Harbor 1989 0000 [0069]
- **H. NEURATH R.L. HILL** *The Proteins* Academic Press 1979 0000 [0076]
- **CUNNINGHAMWELL** *Science*, 1989, vol. 244, 1081-1085 [0078]
- **HILTON et al.** *J. Biol. Chem.*, 1996, vol. 271, 4699-4708 [0078]
- **DE VOS et al.** *Science*, 1992, vol. 255, 306-312 [0078]
- **SMITH et al.** *J. Mol. Biol.*, 1992, vol. 224, 899-904 [0078]
- **WLODAVER et al.** *FEBS Lett.*, 1992, vol. 309, 59-64 [0078]
- **REIDHAAR-OLSONSAUERS** *Science*, 1988, vol. 241, 53-57 [0079]
- **BOWIESAUER** *Proc. Natl. Acad. Sci. USA*, 1989, vol. 86, 2152-2156 [0079]
- **LOWMAN et al.** *Biochemistry*, 1991, vol. 30, 10832-10837 [0079]
- **DERBYSHIRE et al.** *Gene*, 1986, vol. 46, 145- [0079]
- **NER et al.** *DNA*, 1988, vol. 7, 127- [0079]
- **NESS et al.** *Nature Biotechnology*, 1999, vol. 17, 893-896 [0080]
- **COOPER et al.** *EMBO J.*, 1993, vol. 12, 2575-2583 [0082]
- **DAWSON et al.** *Science*, 1994, vol. 266, 776-779 [0082]
- **MARTIN et al.** *J. Ind. Microbiol. Biotechnol.*, 2003, vol. 3, 568-576 [0083]
- **SVETINA et al.** *J. Biotechnol.*, vol. 76, 245-251 [0083]
- **RASMUSSEN-WILSON et al.** *Appl. Environ. Microbiol.*, 1997, vol. 63, 3488-3493 [0083]
- **WARD et al.** *Biotechnology*, 1995, vol. 13, 498-503 [0083]
- **CONTRERAS et al.** *Biotechnology*, 1991, vol. 9, 378-381 [0083]
- **EATON et al.** *Biochemistry*, 1986, vol. 25, 505-512 [0083]
- **COLLINS-RACIE et al.** *Biotechnology*, 1995, vol. 13, 982-987 [0083]
- **CARTER et al.** *Proteins: Structure, Function, and Genetics*, 1989, vol. 6, 240-248 [0083]
- **STEVENS** *Drug Discovery World*, 2003, vol. 4, 35-48 [0083]
- **INNIS et al.** *PCR: A Guide to Methods and Application* Academic Press 1990 0000 [0103]
- **FORD et al.** *Protein Expression and Purification*, 1991, vol. 2, 95-107 [0104]
- **AGAISSELERECLUS** *Molecular Microbiology*, 1994, vol. 13, 97-107 [0108]
- **EGON et al.** *Gene*, 1988, vol. 69, 301-315 [0108]
- **VILLA-KAMAROFF et al.** *Proc. Natl. Acad. Sci. USA*, 1978, vol. 75, 3727-3731 [0108]
- **DEBOER et al.** *Proc. Natl. Acad. Sci. USA*, 1983, vol. 80, 21-25 [0108]
- **GILBERT et al.** *Scientific American*, 1980, vol. 242, 74-94 [0108]
- **ROMANOS et al.** *Yeast*, 1992, vol. 8, 423-488 [0110]
- **HUE et al.** *Journal of Bacteriology*, 1995, vol. 177, 3465-3471 [0116]
- **GUOSHHERMAN** *Mol. Cellular Biol.*, 1995, vol. 15, 5983-5990 [0122]
- **SIMONENPALVA** *Microbiological Reviews*, 1993, vol. 57, 109-137 [0124]
- **GEMS et al.** *Gene*, 1991, vol. 98, 61-67 [0140]
- **CULLEN et al.** *Nucleic Acids Res.*, 1987, vol. 15, 9163-9175 [0140]

- **CHANGCOHEN** Mol. Gen. Genet., 1979, vol. 168, 111-115 [\[0149\]](#)
- **YOUNGSPIZIZEN** J. Bacteriol., 1961, vol. 81, 823-829 [\[0149\]](#)
- **DUBNAUDAVIDOFF-ABELSON** J. Mol. Biol., 1971, vol. 56, 209-221 [\[0149\]](#)
- **SHIGEKAWADOWER** Biotechniques, 1988, vol. 6, 742-751 [\[0149\]](#)
- **KOEHLERTHORNE** J. Bacteriol., 1987, vol. 169, 5271-5278 [\[0149\]](#)
- **HANAHAN** J. Mol. Biol., 1983, vol. 166, 557-580 [\[0149\]](#)
- **DOWER et al.** Nucleic Acids Res., 1988, vol. 16, 6127-6145 [\[0149\]](#)
- **GONG et al.** Folia Microbiol. (Praha), 2004, vol. 49, 399-405 [\[0149\]](#)
- **MAZODIER et al.** J. Bacteriol., 1989, vol. 171, 3583-3585 [\[0149\]](#)
- **BURKE et al.** Proc. Natl. Acad. Sci. USA, 2001, vol. 98, 6289-6294 [\[0149\]](#)
- **CHOI et al.** J. Microbiol. Methods, 2006, vol. 64, 391-397 [\[0149\]](#)
- **PINEDOSMETS** Appl. Environ. Microbiol., 2005, vol. 71, 51-57 [\[0149\]](#)
- **PERRYKURAMITSU** Infect. Immun., 1981, vol. 32, 1295-1297 [\[0149\]](#)
- **CATTJOLICK** Microbios, 1991, vol. 68, 189-207 [\[0149\]](#)
- **BUCKLEY et al.** Appl. Environ. Microbiol., 1999, vol. 65, 3800-3804 [\[0149\]](#)
- **CLEWELL** Microbiol. Rev., 1981, vol. 45, 409-436 [\[0149\]](#)
- **HAWKSWORTH et al.** Ainsworth and Bisby's Dictionary of The Fungi CAB International, University Press 1995 0000 [\[0151\]](#)
- Biology and Activities of Yeast Soc. App. Bacteriol. Symposium Series No. 919800000 [\[0152\]](#)
- **YELTON et al.** Proc. Natl. Acad. Sci. USA, 1984, vol. 81, 1470-1474 [\[0157\]](#)
- **CHRISTENSEN et al.** Technology, 1988, vol. 6, 1419-1422 [\[0157\]](#)
- **MALARDIER et al.** Gene, 1989, vol. 78, 147-156 [\[0157\]](#)
- Guide to Yeast Genetics and Molecular Biology **BECKERGUARENTE** Methods in Enzymology Academic Press, Inc. vol. 194, 182-187 [\[0157\]](#)
- **ITO et al.** J. Bacteriol., 1983, vol. 153, 163- [\[0157\]](#)
- **HINNEN et al.** Proc. Natl. Acad. Sci. USA, 1978, vol. 75, 1920- [\[0157\]](#)
- Protein Purification VCH Publishers 1989 0000 [\[0163\]](#)
- **TAGUE et al.** Plant Physiology, 1988, vol. 86, 506- [\[0172\]](#)
- **FRANCK et al.** Cell, 1980, vol. 21, 285-294 [\[0173\]](#)
- **CHRISTENSEN et al.** Plant Mol. Biol., 1992, vol. 18, 675-689 [\[0173\]](#)
- **ZHANG et al.** Plant Cell, 1991, vol. 3, 1155-1165 [\[0173\]](#)
- **EDWARDSCORUZZI** Ann. Rev. Genet., 1990, vol. 24, 275-303 [\[0173\]](#)
- **ITO et al.** Plant Mol. Biol., 1994, vol. 24, 863-878 [\[0173\]](#)
- **WU et al.** Plant Cell Physiol., 1998, vol. 39, 885-889 [\[0173\]](#)
- **CONRAD et al.** J. Plant Physiol., 1998, vol. 152, 708-711 [\[0173\]](#)
- **CHEN et al.** Plant Cell Physiol., 1998, vol. 39, 935-941 [\[0173\]](#)
- **KYOZUKA et al.** Plant Physiol., 1993, vol. 102, 991-1000 [\[0173\]](#)
- **MITRAHIGGINS** Plant Mol. Biol., 1994, vol. 26, 85-93 [\[0173\]](#)
- **KAGAYA et al.** Mol. Gen. Genet., 1995, vol. 248, 668-674 [\[0173\]](#)
- **XU et al.** Plant Mol. Biol., 1993, vol. 22, 573-588 [\[0173\]](#)
- **GASSER et al.** Science, 1990, vol. 244, 1293- [\[0176\]](#)
- **POTRYKUS** Bio/Technology, 1990, vol. 8, 535- [\[0176\]](#)
- **SHIMAMOTO et al.** Nature, 1989, vol. 338, 274- [\[0176\]](#)

- **HOOYKASSCHILPEROORT** Plant Mol. Biol., 1992, vol. 19, 15-38 [0177]
- **CHRISTOU** Plant J., 1992, vol. 2, 275-281 [0177]
- **SHIMAMOTO** Curr. Opin. Biotechnol., 1994, vol. 5, 158-162 [0177]
- **VASIL et al.** Bio/Technology, 1992, vol. 10, 667-674 [0177]
- **OMIRULLEH et al.** Plant Mol. Biol., 1993, vol. 21, 415-428 [0177]
- Cellulose bioconversion technology **PHILIPPIDIS, G. P.** Handbook on Bioethanol: Production and Utilization Taylor & Francis 1996 0000179-212 [0201]
- **SHEEHAN, J. HIMMEL, M.** Enzymes, energy and the environment: A strategic perspective on the U.S. Department of Energy's research and development activities for bioethanol Biotechnol. Prog., 1999, vol. 15, 817-827 [0201]
- **LYND, L. R. WEIMER, P. J. VAN ZYL, W. H. PRETORIUS, I. S.** Microbial cellulose utilization: Fundamentals and biotechnology Microbiol. Mol. Biol. Reviews, 2002, vol. 66, 506-577 [0201]
- **FERNANDA DE CASTILHOS CORAZZA FLÁVIO FARIA DE MORAES GISELLA MARIA ZANINIVO NEITZEL** Optimal control in fed-batch reactor for the cellobiose hydrolysis Acta Scientiarum. Technology, 2003, vol. 25, 33-38 [0202]
- **GUSAKOV, A. V. SINITSYN, A. P.** Kinetics of the enzymatic hydrolysis of cellulose: 1. A mathematical model for a batch reactor process Enz. Microb. Technol., 1985, vol. 7, 346-352 [0202]
- **RYU, S. K. LEE, J. M.** Bioconversion of waste cellulose by using an attrition bioreactor Biotechnol. Bioeng., 1983, vol. 25, 53-65 [0202]
- **GUSAKOV, A. V. SINITSYN, A. P. DAVYDKIN, I. Y. DAVYDKIN, V. Y. PROTAS, O. V.** Enhancement of enzymatic cellulose hydrolysis using a novel type of bioreactor with intensive stirring induced by electromagnetic field Appl. Biochem. Biotechnol., 1996, vol. 56, 141-153 [0202]
- **CHANDRA et al.** Substrate pretreatment: The key to effective enzymatic hydrolysis of lignocellulosics? Adv. Biochem. Engin./Biotechnol., 2007, vol. 108, 67-93 [0203]
- **GALBEZACCHI** Pretreatment of lignocellulosic materials for efficient bioethanol production Adv. Biochem. Engin./Biotechnol., 2007, vol. 108, 41-65 [0203]
- **HENDRIKS ZEEMAN** Pretreatments to enhance the digestibility of lignocellulosic biomass Bioresource Technol., 2009, vol. 100, 10-18 [0203]
- **MOSIER et al.** Features of promising technologies for pretreatment of lignocellulosic biomass Bioresource Technol., 2005, vol. 96, 673-686 [0203]
- **TAHERZADEH KARIMI** Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: A review Int. J. of Mol. Sci., 2008, vol. 9, 1621-1651 [0203]
- **YANG WYMAN** Pretreatment: the key to unlocking low-cost cellulosic ethanol Biofuels Bioproducts and Biorefining-Biofpr., 2008, vol. 2, 26-40 [0203]
- **DUFF MURRAY** Bioresource Technology, 1996, vol. 855, 1-33 [0207]
- **GALBEZACCHI** Appl. Microbiol. Biotechnol., 2002, vol. 59, 618-628 [0207]
- **BALLESTEROS et al.** Appl. Biochem. Biotechnol., 2006, vol. 129-132, 496-508 [0209]
- **VARGA et al.** Appl. Biochem. Biotechnol., 2004, vol. 113-116, 509-523 [0209]
- **SASSNER et al.** Enzyme Microb. Technol., 2006, vol. 39, 756-762 [0209]
- **SHELL et al.** Bioresource Technol., 2004, vol. 91, 179-188 [0209]
- **LEE et al.** Adv. Biochem. Eng. Biotechnol., 1999, vol. 65, 93-115 [0209]

- **WYMAN et al.**Bioresource Technol., 2005, vol. 96, 1959-1966 [0211]
- **MOSIER et al.**Bioresource Technol., 2005, vol. 96, 673-686 [0211]
- **SCHMIDTTHOMSEN**Bioresource Technol., 1998, vol. 64, 139-151 [0212]
- **PALONEN et al.**Appl. Biochem. Biotechnol., 2004, vol. 117, 1-17 [0212]
- **VARGA et al.**Biotechnol. Bioeng., 2004, vol. 88, 567-574 [0212]
- **MARTIN et al.**J. Chem. Technol. Biotechnol., 2006, vol. 81, 1669-1677 [0212]
- **GOLLAPALLI et al.**Appl. Biochem. Biotechnol., 2002, vol. 98, 23-35 [0214]
- **CHUNDAWAT et al.**Biotechnol. Bioeng., 2007, vol. 96, 219-231 [0214]
- **ALIZADEH et al.**Appl. Biochem. Biotechnol., 2005, vol. 121, 1133-1141 [0214]
- **TEYMOURI et al.**Bioresource Technol., 2005, vol. 96, 2014-2018 [0214]
- **PAN et al.**Biotechnol. Bioeng., 2005, vol. 90, 473-481 [0215]
- **PAN et al.**Biotechnol. Bioeng., 2006, vol. 94, 851-861 [0215]
- **KURABI et al.**Appl. Biochem. Biotechnol., 2005, vol. 121, 219-230 [0215]
- **SCHELL et al.**Appl. Biochem. and Biotechnol., 2003, vol. 105, 10869-85 [0216]
- **MOSIER et al.**Bioresource Technology, 2005, vol. 96, 673-686 [0216]
- Pretreatment of biomass**HSU, T.-A.**Handbook on Bioethanol: Production and UtilizationTaylor& Francis19960000179-212 [0222]
- **GHOSH**SINGHPhysicochemical and biological treatments for enzymatic/microbial conversion of cellulosic biomassAdv. Appl. Microbiol., 1993, vol. 39, 295-333 [0222]
- Pretreating lignocellulosic biomass: a review, in Enzymatic Conversion of Biomass for Fuels Production**MCMILLAN, J. D.**ACS Symposium Series 566American Chemical Society19940000 [0222]
- Ethanol production from renewable resources**GONG, C. S.CAO, N. J.DU, J.TSAO, G. T.**Advances in Biochemical Engineering/BiotechnologySpringer-Verlag19990000vol. 65, 207-241 [0222] [0303]
- **OLSSONHAHN-HAGERDAL**Fermentation of lignocellulosic hydrolysates for ethanol productionEnz. Microb. Tech., 1996, vol. 18, 312-331 [0222]
- **VALLANDERERIKSSON**Production of ethanol from lignocellulosic materials: State of the artAdv. Biochem. Eng./Biotechnol., 1990, vol. 42, 63-95 [0222]
- **PENTTILA et al.**Gene, vol. 45, 253-263 [0250]
- **SALOHEIMO et al.**Gene, 1988, vol. 63, 11-22 [0250]
- **OKADA et al.**Appl. Environ. Microbiol., 1988, vol. 64, 555-563 [0250]
- **SALOHEIMO et al.**Molecular Microbiology, 1994, vol. 13, 219-228 [0250]
- **OOI et al.**Nucleic Acids Research, 1990, vol. 18, 5884- [0250]
- **SAKAMOTO et al.**Current Genetics, 1995, vol. 27, 435-439 [0250]
- **SAARILAHTI et al.**Gene, 1990, vol. 90, 9-14 [0250]
- **KAWAGUCHI et al.**Gene, 1996, vol. 173, 287-288 [0252]
- **DAN et al.**J. Biol. Chem., 2000, vol. 275, 4973-4980 [0252]
- More Gene Manipulations in FungiAcademic Press19910000 [0280]
- **BAILEY, J.E.OLLIS, D.F.**Biochemical Engineering FundamentalsMcGraw-Hill Book Company19860000 [0280]
- **LIN et al.**Appl. Microbiol. Biotechnol., 2006, vol. 69, 627-642 [0287]
- **CHENHO**Cloning and improving the expression of Pichia stipitis xylose reductase gene in Saccharomyces cerevisiaeAppl. Biochem. Biotechnol., 1993, vol. 39-40, 135-

147 [0296]

- **HO et al.** Genetically engineered *Saccharomyces* yeast capable of effectively cofermenting glucose and xylose *Appl. Environ. Microbiol.*, 1998, vol. 64, 1852-1859 [0296]
- **KOTTERCIRIACY** Xylose fermentation by *Saccharomyces cerevisiae* *Appl. Microbiol. Biotechnol.*, 1993, vol. 38, 776-783 [0296]
- **WALFRIDSSON et al.** Xylose-metabolizing *Saccharomyces cerevisiae* strains overexpressing the TKL1 and TAL1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase *Appl. Environ. Microbiol.*, 1995, vol. 61, 4184-4190 [0296]
- **KUYPER et al.** Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle *FEMS Yeast Research*, 2004, vol. 4, 655-664 [0296]
- **BEALL et al.** Parametric studies of ethanol production from xylose and other sugars by recombinant *Escherichia coli*, *Biotech. Bioeng.*, 1991, vol. 38, 296-303 [0296]
- **INGRAM et al.** Metabolic engineering of bacteria for ethanol production *Biotechnol. Bioeng.*, 1998, vol. 58, 204-214 [0296]
- **ZHANG et al.** Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis* *Science*, 1995, vol. 267, 240-243 [0296]
- **DEANDA et al.** Development of an arabinose-fermenting *Zymomonas mobilis* strain by metabolic pathway engineering *Appl. Environ. Microbiol.*, 1996, vol. 62, 4465-4470 [0296]
- The Alcohol Textbook Nottingham University Press 1999 0000 [0300]
- **ALFENORE et al.** Improving ethanol production and viability of *Saccharomyces cerevisiae* by a vitamin feeding strategy during fed-batch process *Springer-Verlag* 2002 0000 [0301]
- **SILVEIRA, M. M. JONAS, R.** The biotechnological production of sorbitol *Appl. Microbiol. Biotechnol.*, 2002, vol. 59, 400-408 [0303]
- **NIGAM, P. SINGH, D.** Processes for fermentative production of xylitol - a sugar substitute *Process Biochemistry*, 1995, vol. 30, 2117-124 [0303]
- **EZEJI, T. C. QURESHI, N. BLASCHEK, H. P.** Production of acetone, butanol and ethanol by *Clostridium beijerinckii* BA101 and in situ recovery by gas stripping *World Journal of Microbiology and Biotechnology*, 2003, vol. 19, 6595-603 [0303]
- **RICHARD, A. MARGARITIS, A.** Empirical modeling of batch fermentation kinetics for poly(glutamic acid) production and other microbial biopolymers *Biotechnology and Bioengineering*, 2004, vol. 87, 4501-515 [0307]
- **KATAOKA, N. A. MIYAK. KIRIYAMA** Studies on hydrogen production by continuous culture system of hydrogen-producing anaerobic bacteria *Water Science and Technology*, 1997, vol. 36, 6-741-47 [0308]
- **GUNASEELAN V. N.** Biomass and Bioenergy, 1997, vol. 13, 1-283-114 [0308]
- **CHEN, R. LEE, Y. Y.** Membrane-mediated extractive fermentation for lactic acid production from cellulosic biomass *Appl. Biochem. Biotechnol.*, 1997, vol. 63-65, 435-448 [0311]
- **STAHLBAND et al.** *J. Biotechnol.*, 1993, vol. 29, 229-242 [0351]

- **KARLSSON et al.** Appl. Microbiol. Biotechnol., 2001, vol. 55, 317-320 [\[0352\]](#)
- **ZHANG** Pulp & Paper Canada, 2000, vol. 101, 59-62 [\[0352\]](#)
- Bacteriological Analytical Manual 1998 0000 [\[0358\]](#) [\[0360\]](#) [\[0361\]](#) [\[0364\]](#)
- **PEARSON et al.** Genomics, 1997, vol. 46, 24-36 [\[0389\]](#)
- **GIESECKE et al.** Journal of Virology Methods, 1992, vol. 38, 47-60 [\[0392\]](#)
- **CHRISTENSEN et al.** Bio/Technology, 1988, vol. 6, 1419-1422 [\[0393\]](#) [\[0406\]](#)
- **RASMUSSEN et al.** Biotechnology and Bioengineering, 2006, vol. 94, 869-876 [\[0416\]](#)

Patentkrav

1. Isoleret polypeptid med endoglucanaseaktivitet, som er valgt fra gruppen bestående af et polypeptid, der udviser mindst 90% sekvensidentitet med aminosyrerne 18 til 419 ifølge SEQ ID NO: 2.
5
2. Polypeptid ifølge krav 1, som omfatter eller består af aminosyrerne 18 til 419 ifølge SEQ ID NO: 2.
- 10 3. Sammensætning, som omfatter polypeptidet ifølge et hvilket som helst af kravene 1-2.
4. Isoleret polynukleotid, som koder for polypeptidet ifølge et hvilket som helst af kravene 1-2.
5. Nukleinsyrekonstruktion eller ekspressionsvektor, som omfatter polynukleotidet ifølge krav 4, hvilket polynukleotid er operativt forbundet til en eller flere kontrolsekvenser, der styrer fremstillingen af polypeptidet i en ekspressionsvært.
15
6. Rekombinant værtscelle, som omfatter polynukleotidet ifølge krav 4, hvilket polynukleotid er operativt forbundet til en eller flere fremmede kontrolsekvenser, der styrer fremstillingen af polypeptidet.
20
7. Fremgangsmåde til fremstilling af et polypeptid med endoglucanaseaktivitet, hvilken fremgangsmåde omfatter:
 - (a) dyrkning af værtscellen ifølge krav 6 under betingelser, som er befordrende for fremstilling af polypeptidet; og
 - (b) indvinding af polypeptidet.
25
8. Transgen plante, plantedel eller plantecelle, som omfatter et polynukleotid, hvilket polynukleotid koder for polypeptidet ifølge et hvilket som helst af kravene 1-2.
30
9. Fremgangsmåde til nedbrydning eller omdannelse af et celluloseholdigt materiale, hvilken fremgangsmåde omfatter: behandling af det celluloseholdige materiale med en enzymsammensætning ifølge krav 3.
- 35 10. Fremgangsmåde ifølge krav 9, hvor enzymsammensætningen omfatter et eller flere enzymer valgt fra gruppen bestående af en cellulase, et GH61-polypeptid med cellulolyseforbedrende aktivitet,

en hemicellulase, et expansin, en esterase, en laccase, et ligninolytisk enzym, en pectinase, en peroxidas og ligninase.

11. Fremgangsmåde til fremstilling af et fermenteringsprodukt, hvilken fremgangsmåde omfatter:

- 5 (a) forsukring af et celluloseholdigt materiale med en enzymsammensætning ifølge krav 3;
- (b) fermentering af det forsukrede celluloseholdige materiale med en eller flere fermenterende mikroorganismer for at fremstille fermenteringsproduktet; og
- (c) indvinding af fermenteringsproduktet fra fermenteringen.

DRAWINGS

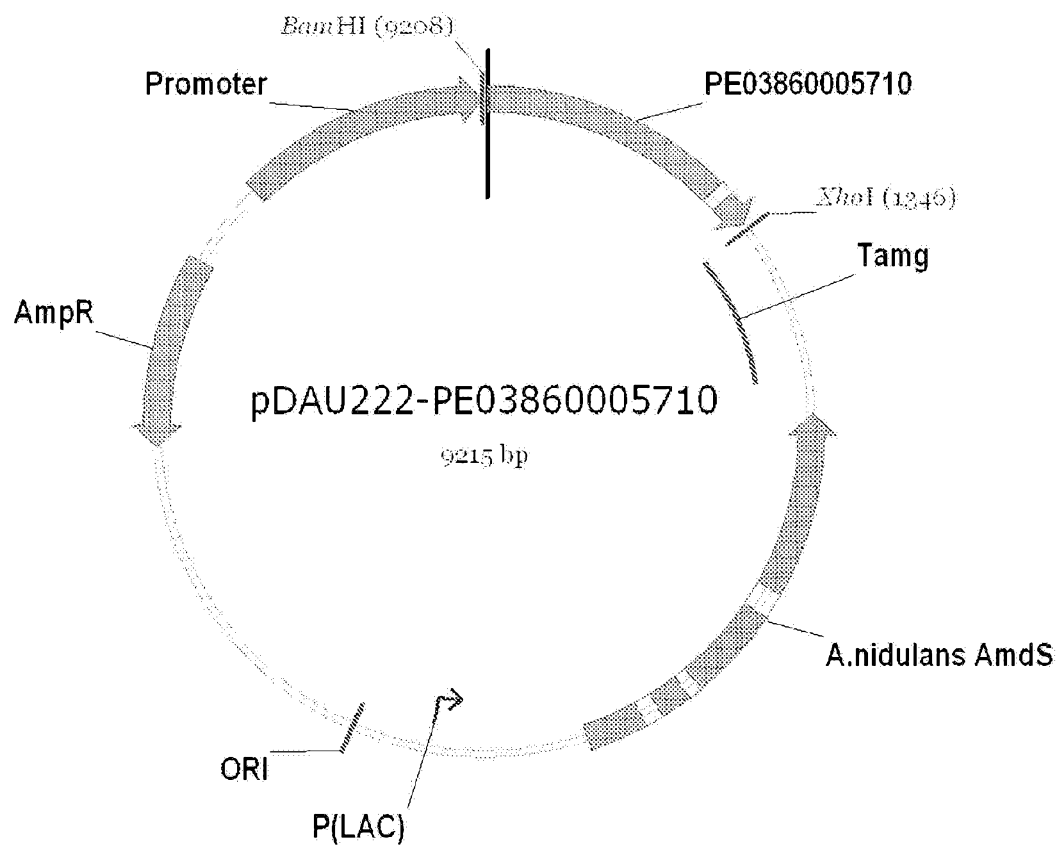
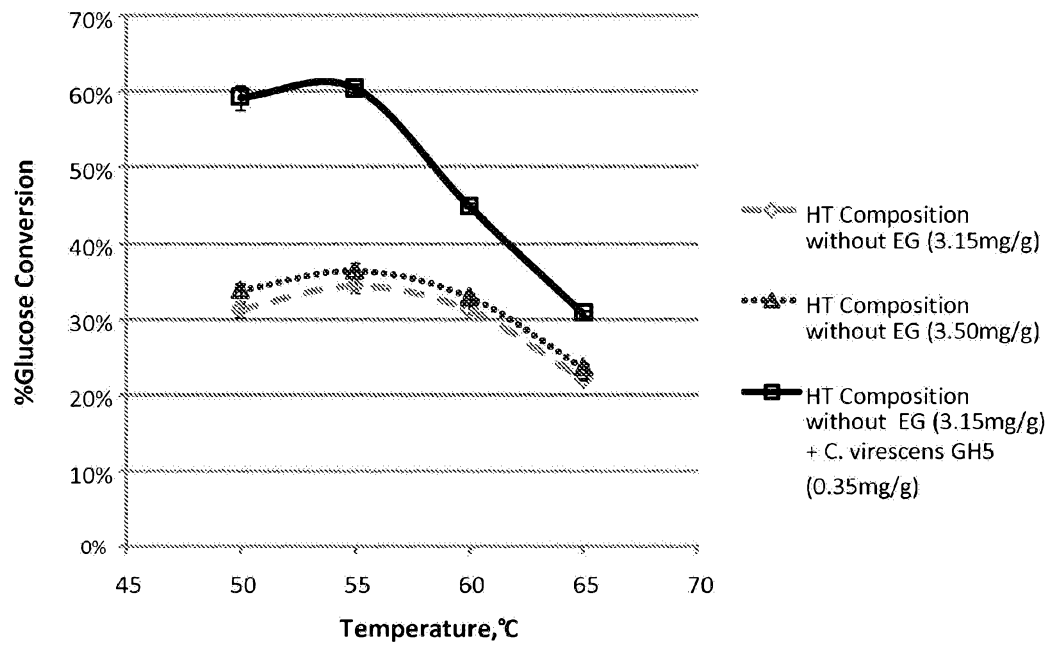
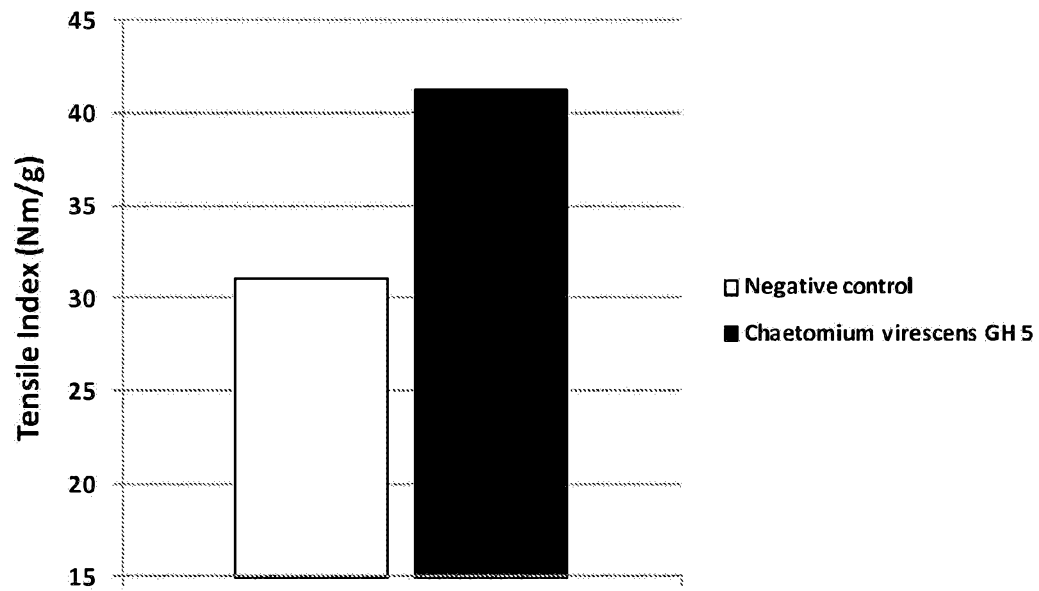
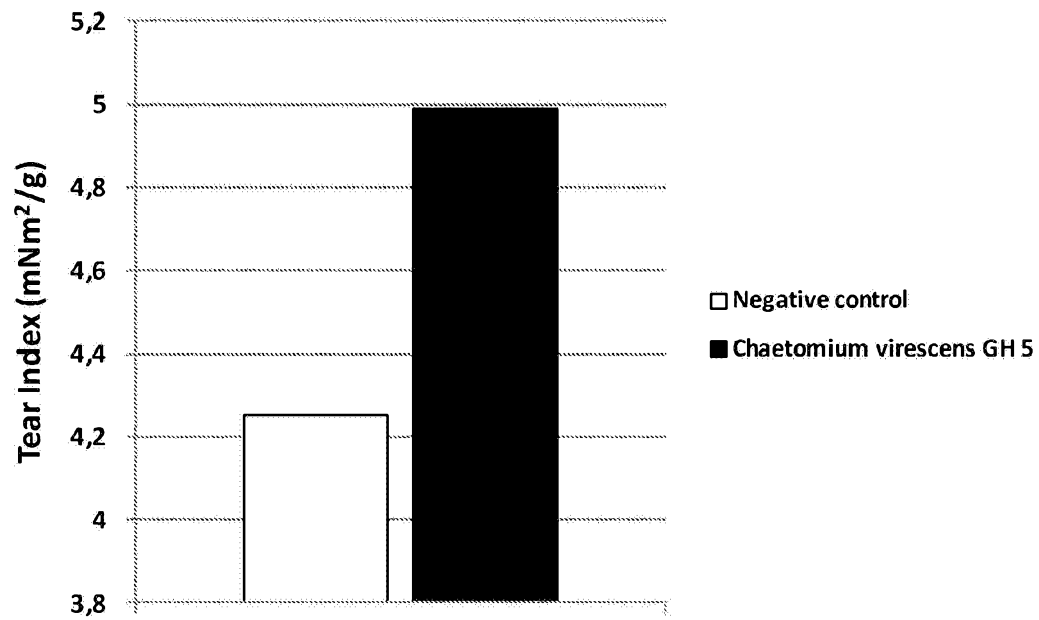
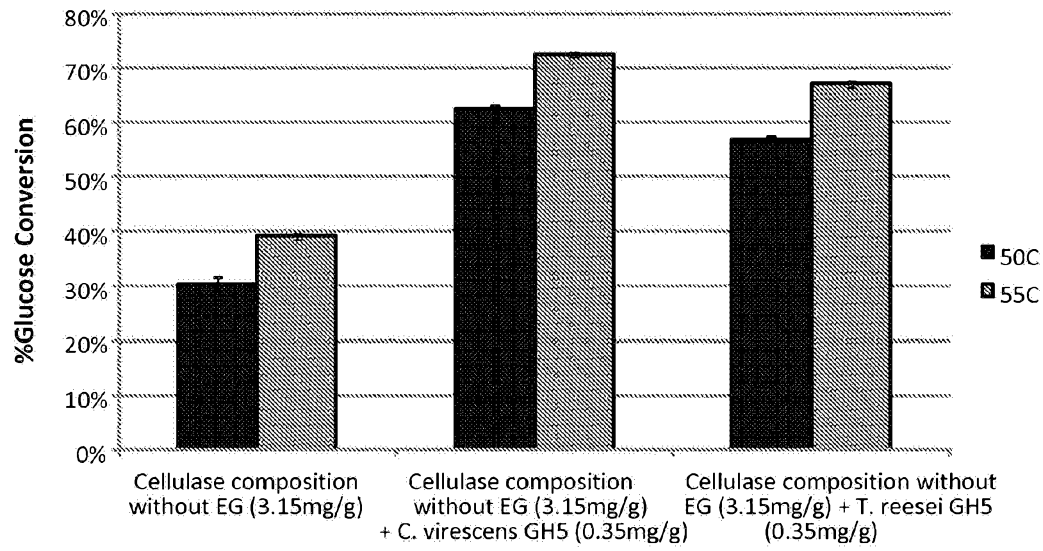


Fig. 1

**Fig. 2**

**Fig. 3**

**Fig. 4**

**Fig. 5**