

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
5 April 2012 (05.04.2012)

(10) International Publication Number
WO 2012/044836 A1

(51) International Patent Classification:

C12N 9/42 (2006.01) D21C 5/00 (2006.01)
C12N 15/82 (2006.01) C12N 5/10 (2006.01)

(21) International Application Number:

PCT/US2011/054036

(22) International Filing Date:

29 September 2011 (29.09.2011)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/388,527 30 September 2010 (30.09.2010) US

(71) Applicant (for all designated States except US):
NOVOZYMES, INC. [US/US]; 1445 Drew Avenue,
Davis, California 95618 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **SWEENEY, Matthew** [US/US]; 517 14th Street, Sacramento, California 95814 (US). **WOGULIS, Mark** [US/US]; 558 Villanova Drive, Davis, California 95616 (US).

(74) Agent: **STARNES, Robert**; Novozymes, Inc., 1445 Drew Avenue, Davis, California 95618 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO,

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

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

Declarations under Rule 4.17:

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

Published:

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

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

— with sequence listing part of description (Rule 5.2(a))

(54) Title: VARIANTS OF POLYPEPTIDES HAVING CELLULOLYTIC ENHANCING ACTIVITY AND POLYNUCLEOTIDES ENCODING SAME

(57) Abstract: The present invention relates to polypeptide having cellulolytic enhancing activity variants. The present invention also relates to polynucleotides encoding the variants; nucleic acid constructs, vectors, and host cells comprising the polynucleotides; and methods of using the variants.



WO 2012/044836 A1

VARIANTS OF POLYPEPTIDES HAVING CELLULOLYTIC ENHANCING ACTIVITY AND POLYNUCLEOTIDES ENCODING SAME

Statement as to Rights to Inventions Made Under Federally Sponsored Research and Development

This invention was made with Government support under Cooperative Agreement DE-FC36-08GO18080 awarded by the Department of Energy. The government has certain rights in this invention.

Cross-Reference to Related Application

This application claims the benefit of U.S. Provisional Application Serial No. 61/388,527 filed September 30, 2010, which is incorporated herein by reference.

Reference to a Sequence Listing

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

Background of the Invention

Field of the Invention

The present invention relates to polypeptide having cellulolytic enhancing activity variants, polynucleotides encoding the variants, methods of producing the variants, and methods of using the variants.

Description of the Related Art

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.

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.

WO 2005/074647, WO 2008/148131, WO 2011/035027 disclose isolated GH61 polypeptides having cellulolytic enhancing activity and the polynucleotides thereof from *Thielavia terrestris*. WO 2005/074656 and WO 2010/065830 disclose isolated GH61
5 polypeptides having cellulolytic enhancing activity and the polynucleotides thereof from *Thermoascus aurantiacus*. WO 2007/089290 discloses an isolated GH61 polypeptide having cellulolytic enhancing activity and the polynucleotide thereof from *Trichoderma reesei*. WO 2009/085935, WO 2009/085859, WO 2009/085864, and WO 2009/085868 disclose isolated GH61 polypeptides having cellulolytic enhancing activity and the polynucleotides thereof
10 from *Myceliophthora thermophila*. WO 2010/138754 discloses isolated GH61 polypeptides having cellulolytic enhancing activity and the polynucleotides thereof from *Aspergillus fumigatus*. WO 2011/005867 discloses isolated GH61 polypeptides having cellulolytic enhancing activity and the polynucleotides thereof from *Penicillium pinophilum*. WO 2011/039319 discloses isolated GH61 polypeptides having cellulolytic enhancing activity and
15 the polynucleotides thereof from *Thermoascus* sp. WO 2011/041397 discloses isolated GH61 polypeptides having cellulolytic enhancing activity and the polynucleotides thereof from *Penicillium* sp. WO 2011/041504 discloses isolated GH61 polypeptides having cellulolytic enhancing activity and the polynucleotides thereof from *Thermoascus crustaceus*. WO 2008/151043 discloses methods of increasing the activity of a GH61
20 polypeptide having cellulolytic enhancing activity by adding a soluble activating divalent metal cation to a composition comprising the polypeptide.

It would be advantageous in the art to improve the ability of polypeptides having cellulolytic enhancing activity to enhance enzymatic degradation of lignocellulosic feedstocks.

25 The present invention provides variants of a polypeptide having cellulolytic enhancing activity with improved properties.

Summary of the Invention

The present invention relates to isolated variants, comprising a substitution at one or
30 more (e.g., several) positions corresponding to positions 75, 77, 179, 181, and 183 of the mature polypeptide of SEQ ID NO: 2, wherein the variants have cellulolytic enhancing activity.

The present invention also relates to isolated polynucleotides encoding the variants; nucleic acid constructs, vectors, and host cells comprising the polynucleotides; and methods
35 of producing the variants.

The present invention also relates to methods for degrading or converting a cellulosic material, comprising: treating the cellulosic material with an enzyme composition in the

presence of such a variant having cellulolytic enhancing activity.

The present invention also relates to methods for producing a fermentation product, comprising:

5 (a) saccharifying a cellulosic material with an enzyme composition in the presence of such a variant having cellulolytic enhancing activity;

(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.

10 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 in the presence of such a variant having cellulolytic enhancing activity.

Brief Description of the Figures

15 Figures 1A and 1B show the genomic DNA sequence and the deduced amino acid sequence of an *Aspergillus fumigatus* gene encoding a GH61B polypeptide having cellulolytic enhancing activity (SEQ ID NOs: 1 and 2, respectively).

20 Figure 2 shows the conversion of phosphoric acid swollen cellulose (0.5% w/w) by the combination of pyrogallol and *A. fumigatus* CEL3A beta-glucosidase; the combination of pyrogallol, *A. fumigatus* GH61B wild-type polypeptide, and *A. fumigatus* CEL3A beta-glucosidase; and the combination of pyrogallol, *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V, and *A. fumigatus* CEL3A beta-glucosidase.

25 Figures 3A and 3B show the effect of the *A. fumigatus* GH61B wild-type polypeptide and *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V on the conversion of pretreated corn stover (PCS) by by the combination of a *Trichoderma reesei* cellulase composition and *A. fumigatus* CEL3A beta-glucosidase at either 50°C or 55°C.

30 Figures 4A and 4B show the effect of the *A. fumigatus* GH61B wild-type polypeptide and *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V on the conversion of PCS by by the combination of a high temperature cellulase composition and *A. fumigatus* CEL3A beta-glucosidase at 50°C, 55°C, 60°C, or 65°C.

Figures 5A and 5B show the T_d (denaturation temperature) of the *Aspergillus fumigatus* wild-type GH61B polypeptide and the *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V by differential scanning calorimetry.

35 Definitions

Acetylxylan esterase: The term “acetylxylan 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, acetylxylan 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 acetylxylan esterase is defined as the amount of enzyme capable of releasing 1 μmole of *p*-nitrophenolate anion per minute at pH 5, 25°C.

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.

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 μl for 30 minutes at 40°C followed by arabinose analysis by AMINEX® HPX-87H column chromatography (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

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 μmole of glucuronic or 4-O-methylglucuronic acid per minute at pH 5, 40°C.

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.

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-(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.

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.

Cellobiohydrolase: The term "cellobiohydrolase" means a 1,4-beta-D-glucan cellobiohydrolase (E.C. 3.2.1.91) that catalyzes the hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, cellobiosaccharides, 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.

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.

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, Wiseloge *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.

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).

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.

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.

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.

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.

5 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.

Cellulolytic enzyme or cellulase: The term "cellulolytic enzyme" or "cellulase" means one or more (e.g., several) enzymes that hydrolyze a cellulosic material. Such
10 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,
15 *Biotechnology Advances* 24: 452-481. Total cellulolytic activity is usually measured using insoluble substrates, including Whatman No1 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 No1 filter paper as the substrate. The assay was established by the International Union of Pure and Applied
20 Chemistry (IUPAC) (Ghose, 1987, Measurement of cellulase activities, *Pure Appl. Chem.* 59: 257-68).

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
25 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).

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

Control sequences: The term "control sequences" means nucleic acid sequences necessary for expression of a polynucleotide encoding a variant of the present invention.

Each control sequence may be native (*i.e.*, from the same gene) or foreign (*i.e.*, from a different gene) to the polynucleotide encoding the variant or native or foreign to each other. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the polynucleotide encoding a variant.

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.

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

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

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.

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.

Fragment: The term "fragment" means a polypeptide having one or more (*e.g.*, several) amino acids absent from the amino and/or carboxyl terminus of a mature polypeptide; wherein the fragment has cellulolytic enhancing activity. In one aspect, a fragment contains at least 200 amino acid residues, *e.g.*, at least 210 amino acid residues or at least 220 amino acid residues of the mature polypeptide of SEQ ID NO: 2.

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 acetylxytan 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.

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.

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.

Increased thermal activity: The term "increased thermal activity" means a higher or broader temperature-dependent activity profile of a variant compared to the temperature-dependent activity profile of the parent. The increased thermal activity of the variant enhances catalysis of a reaction at one or more (*e.g.*, several) specific temperatures relative to the parent. A more thermoactive variant will lead to a decrease in the time required and/or a decrease in the enzyme concentration required for catalysis of the reaction. The increased thermal activity of the variant relative to the parent can be assessed, for example, under conditions of one or more (*e.g.*, several) temperatures. For example, the one or more (*e.g.*, several) temperatures can be any temperature in the range of 25°C to 95°C, *e.g.*, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, or 95°C (or in between) at a pH in the range of 3 to 8, *e.g.*, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, or 8.0, (or in between).

In one aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 40°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 45°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 50°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 55°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 60°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 65°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 70°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 75°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 80°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 85°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.0 and 90°C.

In another aspect, the thermal activity of the variant relative to the parent is determined at pH 3.5 and 40°C. In another aspect, the thermal activity of the variant relative

In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 40°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 45°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 50°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 55°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 60°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 65°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 70°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 75°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 80°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 85°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 7.5 and 90°C.

In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 40°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 45°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 50°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 55°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 60°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 65°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 70°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 75°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 80°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 85°C. In another aspect, the thermal activity of the variant relative to the parent is determined at pH 8.0 and 90°C.

The increased thermal activity of the variant relative to the parent can be determined using any enzyme assay known in the art for GH61 polypeptides having cellulolytic enhancing activity. See for example, WO 2005/074647, WO 2008/148131 WO 2005/074656, WO 2010/065830, WO 2007/089290, WO 2009/085935, WO 2009/085859, WO 2009/085864, WO 2009/085868, and WO 2008/151043, which are incorporated herein by reference. Alternatively, the increased thermal activity of the variant relative to the parent can be determined using any application assay for the variant where the performance of the variant is compared to the parent. For example, the application assay described in Example 12 or 14 can be used.

In one aspect, the thermal activity of the variant is at least 1.01-fold, *e.g.*, at least 1.05-fold, at least 1.1-fold, at least 1.5-fold, at least 1.8-fold, at least 2-fold, at least 5-fold, at least 10-fold, at least 15-fold, at least 20-fold, at least 25-fold, and at least 50-fold more thermally active than the parent.

5 A variant with increased thermal activity may or may not display increased thermostability relative to the parent. For example, a variant may have an improved thermal activity relative to the parent, but does not have increased thermostability.

Increased thermostability: The term "increased thermostability" means a higher retention of cellulolytic enhancing activity of a variant after a period of incubation at a temperature relative to the parent. The increased thermostability of the variant relative to the parent can be assessed, for example, under conditions of one or more (*e.g.*, several) temperatures. For example, the one or more (*e.g.*, several) temperatures can be any temperature in the range of 45°C to 95°C, *e.g.*, 45, 50, 55, 60, 65, 70, 75, 80, 85, or 95°C (or in between) at a pH in the range of 3 to 8, *e.g.*, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 10 or 8.0, (or in between) for a suitable period of incubation, *e.g.*, 1 minute, 5 minutes, 10 minutes, 15 minutes, 30 minutes, 45 minutes, or 60 minutes, such that the variant retains residual activity.

In one aspect, the thermostability of the variant relative to the parent is determined at pH 3.0 and 50°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.0 and 55°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.0 and 60°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.0 and 65°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.0 and 70°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.0 and 75°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.0 and 80°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.0 and 85°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.0 and 90°C.

In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.5 and 50°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.5 and 55°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.5 and 60°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.5 and 65°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.5 and 70°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.5 and 75°C. In another aspect, the thermostability of the variant relative to the parent is determined at pH 3.5 and 80°C. In another aspect, the thermostability of the

incubating the variant and parent for 45 minutes. In each of the aspects above, the thermostability of the variant relative to the parent can be determined by incubating the variant and parent for 60 minutes.

5 The increased thermostability of the variant relative to the parent can be determined by differential scanning calorimetry (DSC) using methods standard in the art (see, for example, Sturtevant, 1987, *Annual Review of Physical Chemistry* 38: 463-488). The increased thermostability of the variant relative to the parent can also be determined using any enzyme assay known in the art for GH61 polypeptides having cellulolytic enhancing activity. See for example, WO 2005/074647, WO 2008/148131 WO 2005/074656, WO 10 2010/065830, WO 2007/089290, WO 2009/085935, WO 2009/085859, WO 2009/085864, WO 2009/085868, and WO 2008/151043, which are incorporated herein by reference. Alternatively, the increased thermostability of the variant relative to the parent can be determined using any application assay for the variant where the performance of the variant is compared to the parent. For example, the application assay described in Example 12 or 15 14 can be used.

In one aspect, the thermostability of the variant having cellulolytic enhancing activity is at least 1.01-fold, e.g., at least 1.05-fold, at least 1.1-fold, at least 1.5-fold, at least 1.8-fold, at least 2-fold, at least 5-fold, at least 10-fold, at least 15-fold, at least 20-fold, at least 25-fold, and at least 50-fold more thermostable than the parent.

20 A variant with increased thermostability may or may not display increased thermal activity relative to the parent. For example, a variant may have an improved ability to refold following incubation at an elevated temperature relative to the parent, but does not have increased thermal activity.

Isolated: The term "isolated" means a substance in a form or environment which 25 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) 30 any substance modified by increasing the amount of the substance relative to other components with which it is naturally associated (e.g., multiple copies of a gene encoding the substance; use of a stronger promoter than the promoter naturally associated with the gene encoding the substance). An isolated substance may be present in a fermentation broth sample.

35 **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.

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 22 to 250 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 21 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.

Mature polypeptide coding sequence: The term "mature polypeptide coding sequence" means a polynucleotide that encodes a mature polypeptide having cellulolytic enhancing activity. In one aspect, the mature polypeptide coding sequence is nucleotides 64 to 859 of SEQ ID NO: 1 based on the SignalP program (Nielsen *et al.*, 1997, *supra*) that predicts nucleotides 1 to 63 of SEQ ID NO: 1 encode a signal peptide. In another aspect, the mature polypeptide coding sequence is the cDNA sequence contained in nucleotides 64 to 859 of SEQ ID NO: 1.

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.

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.

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

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.

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.

Parent or parent polypeptide having cellulolytic enhancing activity: The term “parent” or “parent polypeptide having cellulolytic enhancing activity” means a polypeptide having cellulolytic enhancing activity to which an alteration is made to produce the variants of the present invention. The parent may be a naturally occurring (wild-type) polypeptide or a variant thereof, or a fragment thereof.

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 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, Bagsværd, 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.

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.

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.

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

For purposes of the present invention, the sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *J. Mol. Biol.* 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*,

2000, *Trends Genet.* 16: 276-277), preferably version 3.0.0, 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:

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

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 3.0.0, 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:

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

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 cellulolytic enhancing activity. In one aspect, a subsequence contains at least 600 nucleotides, *e.g.*, at least 630 nucleotides or at least 660 nucleotides of the mature polypeptide coding sequence of SEQ ID NO: 1.

Variants: The term "variant" means a polypeptide having cellulolytic enhancing 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. The variants of the present invention have at least 20%, *e.g.*, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the cellulolytic enhancing activity of the mature polypeptide of SEQ ID NO: 2.

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.

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.

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.

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

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.

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.

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.

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.

Wild-type polypeptide: The term "wild-type polypeptide" means a polypeptide having cellulolytic enhancing activity expressed by a naturally occurring microorganism, such as a bacterium, yeast, or filamentous fungus found in nature.

Detailed Description of the Invention

The present invention relates to isolated variants, comprising a substitution at one or more (e.g., several) positions corresponding to positions 75, 77, 179, 181, and 183 of the mature polypeptide of SEQ ID NO: 2, wherein the variants have cellulolytic enhancing activity.

Conventions for Designation of Variants

For purposes of the present invention, the mature polypeptide disclosed in SEQ ID NO: 2 is used to determine the corresponding amino acid residue in another polypeptide having cellulolytic enhancing activity. The amino acid sequence of another polypeptide having cellulolytic enhancing activity is aligned with the mature polypeptide disclosed in SEQ ID NO: 2, and based on the alignment, the amino acid position number corresponding to any amino acid residue in the mature polypeptide disclosed in SEQ ID NO: 2 is determined using

the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *J. Mol. Biol.* 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *Trends Genet.* 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix.

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

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

For proteins of known structure, several tools and resources are available for retrieving and generating structural alignments. For example, the SCOP superfamilies of proteins have been structurally aligned, and those alignments are accessible and

downloadable. Two or more protein structures can be aligned using a variety of algorithms such as the distance alignment matrix (Holm and Sander, 1998, *Proteins* 33: 88-96) or combinatorial extension (Shindyalov and Bourne, 1998, *Protein Engineering* 11: 739-747), and implementation of these algorithms can additionally be utilized to query structure
5 databases with a structure of interest in order to discover possible structural homologs (e.g., Holm and Park, 2000, *Bioinformatics* 16: 566-567).

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

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

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

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

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

<u>Parent:</u>	<u>Variant:</u>
195	195 195a 195b
G	G - K - A

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

Different alterations. Where different alterations can be introduced at a position, the different alterations are separated by a comma, *e.g.*, “Arg170Tyr,Glu” represents a substitution of arginine at position 170 with tyrosine or glutamic acid. Thus, “Tyr167Gly,Ala + Arg170Gly,Ala” designates the following variants:

- 5 “Tyr167Gly+Arg170Gly”, “Tyr167Gly+Arg170Ala”, “Tyr167Ala+Arg170Gly”, and “Tyr167Ala+Arg170Ala”.

Variants

10 The present invention provides variants comprising a substitution at one or more (*e.g.*, several) positions corresponding to positions 75, 77, 179, 181, and 183 of the mature polypeptide of SEQ ID NO: 2, wherein the variants have cellulolytic enhancing activity.

In an embodiment, the variant has sequence identity of at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, 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%,
15 at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%, but less than 100%, to the amino acid sequence of the parent polypeptide having cellulolytic enhancing activity.

In another embodiment, the variant has at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, at least 80%, at least 80%, at least 81%, at least 82%, at least 83%, at
20 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%, such as at least 96%, at least 97%, at least 98%, or at least 99%, but less than 100%, sequence identity to the mature polypeptide of SEQ ID NO: 2.

In one aspect, the number of substitutions in the variants of the present invention is
25 1-5, such as 1, 2, 3, 4, or 5 substitutions.

In another aspect, a variant comprises a substitution at one or more (*e.g.*, several) positions corresponding to positions 75, 77, 179, 181, and 183. In another aspect, a variant comprises a substitution at two positions corresponding to any of positions 75, 77, 179, 181, and 183. In another aspect, a variant comprises a substitution at three positions
30 corresponding to any of positions 75, 77, 179, 181, and 183. In another aspect, a variant comprises a substitution at four positions corresponding to any of positions 75, 77, 179, 181, and 183. In another aspect, a variant comprises a substitution at each position corresponding to positions 75, 77, 179, 181, and 183.

In another aspect, the variant comprises or consists of a substitution at a position
35 corresponding to position 75. In another aspect, the amino acid at a position corresponding to position 75 is substituted with Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys,

Met, Phe, Pro, Ser, Thr, Trp, Tyr, or Val, preferably with Val. In another aspect, the variant comprises or consists of the substitution I75V of the mature polypeptide of SEQ ID NO: 2.

5 In another aspect, the variant comprises or consists of a substitution at a position corresponding to position 77. In another aspect, the amino acid at a position corresponding to position 77 is substituted with Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, or Val, preferably with Leu. In another aspect, the variant comprises or consists of the substitution F77L of the mature polypeptide of SEQ ID NO: 2.

10 In another aspect, the variant comprises or consists of a substitution at a position corresponding to position 179. In another aspect, the amino acid at a position corresponding to position 179 is substituted with Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, or Val, preferably with Ile. In another aspect, the variant comprises or consists of the substitution F179I of the mature polypeptide of SEQ ID NO: 2.

15 In another aspect, the variant comprises or consists of a substitution at a position corresponding to position 181. In another aspect, the amino acid at a position corresponding to position 181 is substituted with Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, or Val, preferably with Leu. In another aspect, the variant comprises or consists of the substitution I181L of the mature polypeptide of SEQ ID NO: 2.

20 In another aspect, the variant comprises or consists of a substitution at a position corresponding to position 183. In another aspect, the amino acid at a position corresponding to position 183 is substituted with Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, or Val, preferably with Val. In another aspect, the variant comprises or consists of the substitution I183V of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75 and 77, such as those described above.

25 In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75 and 179, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75 and 181, such as those described above.

30 In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75 and 183, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 77 and 179, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 77 and 181, such as those described above.

35 In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 77 and 183, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 179 and 181, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 179 and 183, such as those described above.

5 In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 181 and 183, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 77, and 179, such as those described above.

10 In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 77, and 181, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 77, and 183, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 179, and 181, such as those described above.

15 In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 179, and 183, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 181, and 183, such as those described above.

20 In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 77, 179, and 181, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 77, 179, and 183, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 179, 181, and 183, such as those described above.

25 In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 77, 181, and 183, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 77, 179, and 181, such as those described above.

30 In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 179, 181, and 183, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 77, 181, and 183, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 77, 179, and 183, such as those described above.

35 In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 77, 179, 181, and 183, such as those described above.

In another aspect, the variant comprises or consists of substitutions at positions corresponding to positions 75, 77, 179, 181, and 183, such as those described above.

In another aspect, the variant comprises or consists of one or more (*e.g.*, several) substitutions selected from the group consisting of I75V, F77L, F179I, I181L, and I183V.

5 In another aspect, the variant comprises or consists of the substitutions I75V + F77L of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions I75V + F179I of the mature polypeptide of SEQ ID NO: 2.

10 In another aspect, the variant comprises or consists of the substitutions I75V + I181L of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions I75V + I183V of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions F77L + F179I of the mature polypeptide of SEQ ID NO: 2.

15 In another aspect, the variant comprises or consists of the substitutions F77L + I181L of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions F77L + I183V of the mature polypeptide of SEQ ID NO: 2.

20 In another aspect, the variant comprises or consists of the substitutions F179I + I181L of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions F179I + I183V of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

25 In another aspect, the variant comprises or consists of the substitutions I75V + F77L + F179I of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions I75V + F77L + I181L of the mature polypeptide of SEQ ID NO: 2.

30 In another aspect, the variant comprises or consists of the substitutions I75V + F77L + I183V of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions I75V + F179I + I181L of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions I75V + F179I + I183V of the mature polypeptide of SEQ ID NO: 2.

35 In another aspect, the variant comprises or consists of the substitutions I75V + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions F77L + F179I + I181L of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions F77L + F179I + I183V of the mature polypeptide of SEQ ID NO: 2.

5 In another aspect, the variant comprises or consists of the substitutions F179I + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions F77L + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

10 In another aspect, the variant comprises or consists of the substitutions I75V + F77L + F179I + I181L of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions I75V + F179I + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions I75V + F77L + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

15 In another aspect, the variant comprises or consists of the substitutions I75V + F77L + F179I + I183V of the mature polypeptide of SEQ ID NO: 2.

In another aspect, the variant comprises or consists of the substitutions F77L + F179I + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

20 In another aspect, the variant comprises or consists of the substitutions I75V + F77L + F179I + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

The variants of the present invention may further comprise an alteration, *e.g.*, deletion, insertion, and/or substitution at one or more (*e.g.*, several) other positions.

25 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.

30 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.
35 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.

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.

5 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 polypeptide having cellulolytic enhancing activity to identify amino
10 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 polypeptide having cellulolytic enhancing activity 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
15 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.

The variants may consist of 200 to 220 amino acids, *e.g.*, 205 to 220, 210 to 220,
20 and 215 to 220 amino acids.

In an embodiment, the variant has increased thermal activity.

In an embodiment, the variant has increased thermostability.

In an embodiment, the variant has increased thermal activity and increased
thermostability.

25

Parent Polypeptides Having Cellulolytic Enhancing Activity

The parent polypeptide having cellulolytic enhancing activity may be (a) a polypeptide having at least 60% sequence identity to the mature polypeptide of SEQ ID NO: 2; (b) a polypeptide encoded by a polynucleotide that hybridizes under at least low
30 stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, (ii) the cDNA sequence thereof, or (iii) the full-length complement of (i) or (ii); or (c) a polypeptide encoded by a polynucleotide having at least 60% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof.

In one aspect, the parent has a sequence identity to the mature polypeptide of SEQ
35 ID NO: 2 of at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, 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%, which have cellulolytic enhancing activity. In one aspect, the amino acid sequence of the parent differs by up to 10 amino acids, e.g., by 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids, from the mature polypeptide of SEQ ID NO: 2.

5 In another aspect, the parent comprises or consists of the amino acid sequence of SEQ ID NO: 2. In another aspect, the parent comprises or consists of the mature polypeptide of SEQ ID NO: 2. In another aspect, the parent comprises or consists of amino acids 22 to 250 of SEQ ID NO: 2.

10 In another aspect, the parent is a fragment of the mature polypeptide of SEQ ID NO: 2 containing at least 200 amino acid residues, e.g., at least 210 amino acid residues or at least 220 amino acid residues.

In another aspect, the parent is an allelic variant of the mature polypeptide of SEQ ID NO: 2.

15 In another aspect, the parent is encoded by a polynucleotide that hybridizes 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) (J. Sambrook, E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, New York).

20 The polynucleotide of SEQ ID NO: 1 or a subsequence thereof, as well as the polypeptide of SEQ ID NO: 2 or a fragment thereof, may be used to design nucleic acid probes to identify and clone DNA encoding a parent from strains of different genera or species according to methods well known in the art. In particular, such probes can be used
25 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
30 nucleotides, at least 400 nucleotides, at least 500 nucleotides, at least 600 nucleotides, at least 700 nucleotides, at least 800 nucleotides, or at least 900 nucleotides in length. Both DNA and RNA probes can be used. The probes are typically labeled for detecting the corresponding gene (for example, with ³²P, ³H, ³⁵S, biotin, or avidin). Such probes are encompassed by the present invention.

35 A genomic DNA or cDNA library prepared from such other strains may be screened for DNA that hybridizes with the probes described above and encodes a parent. Genomic or other DNA from such other strains may be separated by agarose or polyacrylamide gel

electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to and immobilized on nitrocellulose or other suitable carrier material. In order to identify a clone or DNA that hybridizes with SEQ ID NO: 1 or a subsequence thereof, the carrier material is used in a Southern blot.

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

In one aspect, the nucleic acid probe is the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof. In another aspect, the nucleic acid probe is nucleotides 64 to 859 of SEQ ID NO: 1 or the cDNA sequence thereof. In another aspect,
15 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.

In another aspect, the parent is encoded by a polynucleotide having a sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof of at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, at least 80%, at least
20 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%, which encodes a polypeptide having cellulolytic enhancing activity. In one aspect, the mature
25 polypeptide coding sequence is nucleotides 64 to 859 of SEQ ID NO: 1.

The parent 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.

The parent may be a fusion polypeptide or cleavable fusion polypeptide in which another polypeptide is fused at the N-terminus or the C-terminus of the polypeptide of the
30 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
35 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).

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.

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

The parent may be a bacterial polypeptide having cellulolytic enhancing activity. For example, the parent may be a Gram-positive bacterial polypeptide such as a *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, or *Streptomyces* polypeptide having cellulolytic enhancing activity, or a Gram-negative bacterial polypeptide such as a *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, or *Ureaplasma* polypeptide having cellulolytic enhancing activity.

In one aspect, the parent 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 cellulolytic enhancing activity.

In another aspect, the parent is a *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, or *Streptococcus equi* subsp. *Zooepidemicus* polypeptide having cellulolytic enhancing activity.

In another aspect, the parent is a *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, or *Streptomyces lividans* polypeptide having cellulolytic enhancing activity.

The parent may be a fungal polypeptide having cellulolytic enhancing activity. For example, the parent may be a yeast polypeptide having cellulolytic enhancing activity such as a *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* polypeptide having cellulolytic enhancing activity. For example, the parent may be a filamentous fungal polypeptide having cellulolytic enhancing activity 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*,
5 *Melanocarpus*, *Meripilus*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*,
Paecilomyces, *Penicillium*, *Phanerochaete*, *Piromyces*, *Poitrasia*, *Pseudoplectania*,
Pseudotriconympha, *Rhizomucor*, *Schizophyllum*, *Scytalidium*, *Talaromyces*,
Thermoascus, *Thielavia*, *Tolypocladium*, *Trichoderma*, *Trichophaea*, *Verticillium*, *Volvariella*,
or *Xylaria* polypeptide having cellulolytic enhancing activity.

10 In another aspect, the parent is a *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, or *Saccharomyces oviformis* polypeptide having cellulolytic enhancing activity.

In another aspect, the parent is an *Acremonium cellulolyticus*, *Aspergillus aculeatus*,
15 *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*,
Aspergillus nidulans, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium inops*,
Chrysosporium keratinophilum, *Chrysosporium lucknowense*, *Chrysosporium merdarium*,
Chrysosporium pannicola, *Chrysosporium queenslandicum*, *Chrysosporium tropicum*,
Chrysosporium zonatum, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium*
20 *crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium*
heterosporum, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium*
roseum, *Fusarium sambucinum*, *Fusarium sarcochromum*, *Fusarium sporotrichioides*,
Fusarium sulphureum, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium*
venenatum, *Humicola grisea*, *Humicola insolens*, *Humicola lanuginosa*, *Irpex lacteus*, *Mucor*
25 *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 setosa*, *Thielavia*
spededonium, *Thielavia subthermophila*, *Thielavia terrestris*, *Trichoderma harzianum*,
30 *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma*
viride polypeptide having cellulolytic enhancing activity.

In another aspect, the parent is an *Aspergillus fumigatus* polypeptide having cellulolytic enhancing activity, e.g., the polypeptide comprising the mature polypeptide of SEQ ID NO: 2.

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

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

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
5 Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau Voor Schimmelcultures (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

The parent may be identified and obtained from other sources including microorganisms isolated from nature (*e.g.*, soil, composts, water, etc.) or DNA samples
10 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 a parent 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 parent has been detected with the probe(s),
15 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*).

Preparation of Variants

The present invention also relates to methods for obtaining a variant having
20 cellulolytic enhancing activity, comprising: (a) introducing into a parent polypeptide having cellulolytic enhancing activity a substitution at one or more (*e.g.*, several) positions corresponding to positions 75, 77, 179, 181, and 183 of the mature polypeptide of SEQ ID NO: 2, wherein the variant has cellulolytic enhancing activity; and (b) recovering the variant.

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

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

Site-directed mutagenesis can be accomplished *in vitro* by PCR involving the use of
30 oligonucleotide primers containing the desired mutation. Site-directed mutagenesis can also be performed *in vitro* by cassette mutagenesis involving the cleavage by a restriction enzyme at a site in the plasmid comprising a polynucleotide encoding the parent and subsequent ligation of an oligonucleotide containing the mutation in the polynucleotide.
35 Usually the restriction enzyme that digests the plasmid and the oligonucleotide is the same, permitting sticky ends of the plasmid and the insert to ligate to one another. See, *e.g.*,

Scherer and Davis, 1979, *Proc. Natl. Acad. Sci. USA* 76: 4949-4955; and Barton *et al.*, 1990, *Nucleic Acids Res.* 18: 7349-4966.

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

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

Synthetic gene construction entails *in vitro* synthesis of a designed polynucleotide molecule to encode a polypeptide of interest. Gene synthesis can be performed utilizing a number of techniques, such as the multiplex microchip-based technology described by Tian *et al.* (2004, *Nature* 432: 1050-1054) and similar technologies wherein oligonucleotides are synthesized and assembled upon photo-programmable microfluidic chips.

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).

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.

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

Polynucleotides

The present invention also relates to isolated polynucleotides encoding a variant of

the present invention.

Nucleic Acid Constructs

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

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

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

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

35 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.

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

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

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

Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* 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.

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

The control sequence may also be an mRNA stabilizer region downstream of a promoter and upstream of the coding sequence of a variant of the present invention which increases expression of the variant.

10 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).

The control sequence may also be a leader, a nontranslated region of an mRNA that is important for translation by the host cell. The leader sequence is operably linked to the
15 5'-terminus of the polynucleotide encoding the variant. Any leader that is functional in the host cell may be used.

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

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

The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3'-terminus of the variant-encoding sequence and, when transcribed,
25 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.

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
30 *Fusarium oxysporum* trypsin-like protease.

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

The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a variant and directs the variant into the cell's
35 secretory pathway. The 5'-end of the coding sequence of the polynucleotide may inherently contain a signal peptide coding sequence naturally linked in translation reading frame with the segment of the coding sequence that encodes the variant. Alternatively, the 5'-end of the

coding sequence may contain a signal peptide coding sequence that is foreign to the coding sequence. A foreign signal peptide coding sequence may be required where the coding sequence does not naturally contain a signal peptide coding sequence. Alternatively, a foreign signal peptide coding sequence may simply replace the natural signal peptide coding sequence in order to enhance secretion of the variant. However, any signal peptide coding sequence that directs the expressed variant into the secretory pathway of a host cell may be used.

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.

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.

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

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

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

It may also be desirable to add regulatory sequences that regulate expression of the variant relative to the growth of the host cell. Examples of regulatory 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* TAKA 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 variant would be operably linked to the regulatory sequence.

Expression Vectors

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

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

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

The vector preferably contains one or more (*e.g.*, several) selectable markers that permit easy selection of transformed, transfected, transduced, or the like cells. A selectable

marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

Examples of bacterial selectable markers are *Bacillus licheniformis* or *Bacillus subtilis* *dal* genes, or markers that confer antibiotic resistance such as ampicillin, chloramphenicol, 5 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* (phosphoribosyl-aminoimidazole synthase), *amdS* (acetamidase), *argB* (ornithine 10 carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hph* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenylyltransferase), and *trpC* (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are *Aspergillus nidulans* or *Aspergillus oryzae* *amdS* and *pyrG* genes and a *Streptomyces hygrosopicus* *bar* gene. Preferred for 15 use in a *Trichoderma* cell are *adeA*, *adeB*, *amdS*, *hph*, and *pyrG* genes.

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.

The vector preferably contains an element(s) that permits integration of the vector 20 into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

For integration into the host cell genome, the vector may rely on the polynucleotide's sequence encoding the variant or any other element of the vector for integration into the genome by homologous or non-homologous recombination. Alternatively, the vector may 25 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 30 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.

35 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*.

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*.

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

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

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

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

The present invention also relates to recombinant host cells, comprising a polynucleotide encoding a variant of the present invention operably linked to one or more (*e.g.*, several) control sequences that direct the production of a variant 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 variant and the source of the parent.

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

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

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

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.

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

The introduction of DNA into a *Bacillus* cell may 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.

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

5 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).

10 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).

15 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*,
20 *Saccharomyces oviformis*, or *Yarrowia lipolytica* cell.

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

The filamentous fungal host cell may be an *Acremonium*, *Aspergillus*,
30 *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.

For example, the filamentous fungal host cell may be an *Aspergillus awamori*,
35 *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*,
5 *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*,
Fusarium graminum, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*,
Fusarium reticulatum, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochromum*,
Fusarium sporotrichioides, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium*
trichothecioides, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor*
10 *miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*,
Phanerochaete chrysosporium, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia terrestris*,
Trametes villosa, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*,
Trichoderma longibrachiatum, *Trichoderma reesei*, or *Trichoderma viride* cell.

Fungal cells may be transformed by a process involving protoplast formation,
15 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
20 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.

25 **Methods of Production**

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

30 The host cells are cultivated in a nutrient medium suitable for production of the
variant 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 performed in a
suitable medium and under conditions allowing the variant to be expressed and/or isolated.
35 The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen
sources and inorganic salts, using procedures known in the art. Suitable media are available
from commercial suppliers or may be prepared according to published compositions (e.g., in

catalogues of the American Type Culture Collection). If the variant is secreted into the nutrient medium, the variant can be recovered directly from the medium. If the variant is not secreted, it can be recovered from cell lysates.

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

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

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

Compositions

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

25 The compositions may comprise a variant 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 (several) enzymes selected from the group consisting of a cellulase, a 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.

30 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.

35 The compositions may be a fermentation broth formulation or a cell composition, as described herein. Consequently, the present invention also relates to fermentation broth formulations and cell compositions comprising a variant of the present invention. In some embodiments, the composition is a cell-killed whole broth containing organic acid(s), killed cells and/or cell debris, and culture medium.

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 embodiments, the fermentation broth contains spent cell culture medium, extracellular enzymes, and viable and/or nonviable microbial cells.

In an embodiment, 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 embodiment, 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.

In one aspect, the composition contains an organic acid(s), and optionally further contains killed cells and/or cell debris. In one embodiment, 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.

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.

The cell-killed whole broth or composition may further comprise one or more enzyme activities such as cellobiohydrolase, endoglucanase, beta-glucosidase, endo-beta-1,3(4)-glucanase, glucohydrolase, xyloglucanase, xylanase, xylosidase, arabinofuranosidase, alpha-glucuronidase, acetyl xylan esterase, mannanase, mannosidase, alpha-galactosidase, mannan acetyl esterase, galactanase, arabinanase, pectate lyase, pectinase lyase, pectate lyase, polygalacturonase, pectin acetyl esterase, pectin methyl esterase, beta-galactosidase, galactanase, arabinanase, alpha-arabinofuranosidase, rhamnogalacturonase, ferrulic acid esterases rhamnogalacturonan lyase, rhamnogalacturonan acetyl esterase, xylogalacturonosidase, xylogalacturonase, rhamnogalacturonan lyase, lignin peroxidases, manganese-dependent peroxidases, hybrid peroxidases, with combined properties of lignin

peroxidases and manganese-dependent peroxidases, glucoamylase, amylase, protease, and laccase.

In some embodiments, the cell-killed whole broth or composition includes cellulolytic enzymes including, but not limited to, (i) endoglucanases (EG) or 1,4-D-glucan-4-
5 glucohydrolases (EC 3.2.1.4), (ii) exoglucanases, including 1,4-D-glucan glucohydrolases (also known as celldextnnases) (EC 3.2.1.74) and 1,4-D-glucan cellobiohydrolases (exo-cellobiohydrolases, CBH) (EC 3.2.1.91), and (iii) beta-glucosidase (BG) or beta-glucoside glucohydrolases (EC 3.2.1.21).

The cell-killed whole broth or composition may contain the unfractionated contents of
10 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 (*e.g.*, expression of cellulase and/or glucosidase enzyme(s)). In some embodiments, the cell-killed whole broth or composition
15 contains the spent cell culture medium, extracellular enzymes, and killed filamentous fungal cells. In some embodiments, the microbial cells present in the cell-killed whole broth or composition can be permeabilized and/or lysed using methods known in the art.

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,
20 and/or insoluble enzyme(s). In some embodiments, insoluble components may be removed to provide a clarified liquid composition.

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.

Examples are given below of preferred uses of the compositions of the present
25 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

The present invention is also directed to the following methods for using the variants,
30 or compositions thereof.

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

The present invention also relates to methods for producing a fermentation product, comprising: (a) saccharifying a cellulosic material with an enzyme composition in the presence of a variant 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.

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 in the presence of a variant 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.

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.

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.

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
5 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
10 hydrolysis (saccharification), fermentation, or a combination thereof, can be used in the practicing the methods of the present invention.

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
15 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.*
20 *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
25 type reactors for hydrolysis and/or fermentation.

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
30 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
35 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).

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

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.

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).

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.

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.

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).

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).

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.

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.

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).

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.

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.

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.

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.

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.

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).

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.

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.

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).

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 variant of the present invention. The components of the compositions can be added simultaneously or sequentially.

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 components, i.e., optimal for the enzyme components. 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.

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 %.

The enzyme compositions can comprise any protein useful in degrading the cellulosic material.

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 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.

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 an endoglucanase, 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.

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).

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.

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

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.

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.

The optimum amounts of the enzymes and a variant of the present invention depend on several factors including, but not limited to, the mixture of cellulolytic and/or hemicellulolytic enzyme components, the cellulosic substrate, the concentration of cellulosic substrate, the pretreatment(s) of the cellulosic substrate, temperature, time, pH, and inclusion of fermenting organism (*e.g.*, yeast for Simultaneous Saccharification and Fermentation).

In one aspect, an effective amount of cellulolytic or hemicellulolytic enzyme protein to cellulosic material is about 0.5 to about 50 mg, preferably at about 0.5 to about 40 mg, more preferably at about 0.5 to about 25 mg, more preferably at about 0.75 to about 20 mg, more preferably at about 0.75 to about 15 mg, even more preferably at about 0.5 to about 10 mg, and most preferably at about 2.5 to about 10 mg per g of cellulosic material.

In another aspect, an effective amount of a variant having cellulolytic enhancing activity to cellulosic material is about 0.01 to about 50.0 mg, preferably about 0.01 to about 40 mg, more preferably about 0.01 to about 30 mg, more preferably about 0.01 to about 20 mg, more preferably about 0.01 to about 10 mg, more preferably about 0.01 to about 5 mg, more preferably at about 0.025 to about 1.5 mg, more preferably at about 0.05 to about 1.25 mg, more preferably at about 0.075 to about 1.25 mg, more preferably at about 0.1 to about 1.25 mg, even more preferably at about 0.15 to about 1.25 mg, and most preferably at about 0.25 to about 1.0 mg per g of cellulosic material.

In another aspect, an effective amount of a variant having cellulolytic enhancing activity to cellulolytic enzyme protein is about 0.005 to about 1.0 g, preferably at about 0.01 to about 1.0 g, more preferably at about 0.15 to about 0.75 g, more preferably at about 0.15 to about 0.5 g, more preferably at about 0.1 to about 0.5 g, even more preferably at about

0.1 to about 0.5 g, and most preferably at about 0.05 to about 0.2 g per g of cellulolytic enzyme protein.

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.

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.

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.

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

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

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*,
5 *Diplodia*, *Exidia*, *Filibasidium*, *Fusarium*, *Gibberella*, *Holomastigotoides*, *Humicola*, *Irpex*, *Lentinula*, *Leptosphaeria*, *Magnaporthe*, *Melanocarpus*, *Meripilus*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Piromyces*, *Poitrasia*, *Pseudoplectania*, *Pseudotriconympha*, *Rhizomucor*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trichoderma*, *Trichophaea*,
10 *Verticillium*, *Volvariella*, or *Xylaria* polypeptide having enzyme activity.

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.

In another aspect, the polypeptide is an *Acremonium cellulolyticus*, *Aspergillus*
15 *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*,
20 *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochromum*, *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*,
30 *Trichoderma viride*, or *Trichophaea saccata* polypeptide having enzyme activity.

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

One or more (e.g., several) components of the enzyme composition may be a
35 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.

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), CELLUCLAST™ (Novozymes A/S), NOVOZYME™ 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 (Röhm 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.

Examples of bacterial endoglucanases that can be used in the methods of the present invention, include, but are not limited to, an *Acidotherrmus 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).

Examples of fungal endoglucanases that can be used in the present invention include, but are not limited to, a *Trichoderma reesei* endoglucanase I (Penttila *et al.*, 1986, *Gene* 45: 253-263; *Trichoderma reesei* Cel7B endoglucanase I; GENBANK™ accession no. M15665; SEQ ID NO: 4); *Trichoderma reesei* endoglucanase II (Saloheimo, *et al.*, 1988, *Gene* 63:11-22; *Trichoderma reesei* Cel5A endoglucanase II; GENBANK™ accession no. M19373; SEQ ID NO: 6); *Trichoderma reesei* endoglucanase III (Okada *et al.*, 1988, *Appl. Environ. Microbiol.* 64: 555-563; GENBANK™ accession no. AB003694; SEQ ID NO: 8); *Trichoderma reesei* endoglucanase V (Saloheimo *et al.*, 1994, *Molecular Microbiology* 13: 219-228; GENBANK™ accession no. Z33381; SEQ ID NO: 10); *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 (SEQ ID NO: 12); *Myceliophthora thermophila* CBS 117.65 endoglucanase (SEQ ID NO: 14); basidiomycete CBS 495.95 endoglucanase (SEQ ID NO: 16); basidiomycete CBS 494.95 endoglucanase (SEQ ID NO: 18); *Thielavia terrestris* NRRL 8126 CEL6B endoglucanase (SEQ ID NO: 20);
5 *Thielavia terrestris* NRRL 8126 CEL6C endoglucanase (SEQ ID NO: 22); *Thielavia terrestris* NRRL 8126 CEL7C endoglucanase (SEQ ID NO: 24); *Thielavia terrestris* NRRL 8126 CEL7E endoglucanase (SEQ ID NO: 26); *Thielavia terrestris* NRRL 8126 CEL7F endoglucanase (SEQ ID NO: 28); *Cladorrhinum foecundissimum* ATCC 62373 CEL7A endoglucanase (SEQ ID NO: 30); and *Trichoderma reesei* strain No. VTT-D-80133
10 endoglucanase (SEQ ID NO: 32; GENBANK™ accession no. M15665). The endoglucanases of SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, and SEQ ID NO: 32, described above are encoded by the mature polypeptide coding sequence of SEQ ID NO: 3, SEQ ID NO: 5, SEQ
15 ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, and SEQ ID NO: 31, respectively.

Examples of cellobiohydrolases useful in the present invention include, but are not limited to, *Trichoderma reesei* cellobiohydrolase I (SEQ ID NO: 34); *Trichoderma reesei*
20 cellobiohydrolase II (SEQ ID NO: 36); *Humicola insolens* cellobiohydrolase I (SEQ ID NO: 38); *Myceliophthora thermophila* cellobiohydrolase II (SEQ ID NO: 40 and SEQ ID NO: 42); *Thielavia terrestris* cellobiohydrolase II (CEL6A) (SEQ ID NO: 44); *Chaetomium thermophilum* cellobiohydrolase I (SEQ ID NO: 46); and *Chaetomium thermophilum* cellobiohydrolase II (SEQ ID NO: 48), *Aspergillus fumigatus* cellobiohydrolase I (SEQ ID NO:
25 50), and *Aspergillus fumigatus* cellobiohydrolase II (SEQ ID NO: 52). The cellobiohydrolases of SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, and SEQ ID NO: 52, described above are encoded by the mature polypeptide coding sequence of SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45,
30 SEQ ID NO: 47, SEQ ID NO: 49, and SEQ ID NO: 51, respectively.

Examples of beta-glucosidases useful in the present invention include, but are not limited to, *Aspergillus oryzae* beta-glucosidase (SEQ ID NO: 54); *Aspergillus fumigatus* beta-glucosidase (SEQ ID NO: 56); *Penicillium brasilianum* IBT 20888 beta-glucosidase (SEQ ID NO: 58); *Aspergillus niger* beta-glucosidase (SEQ ID NO: 60); and *Aspergillus aculeatus*
35 beta-glucosidase (SEQ ID NO: 62). The beta-glucosidases of SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, and SEQ ID NO: 61 described above are encoded by the mature polypeptide coding sequence of SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO:

57, SEQ ID NO: 59, and SEQ ID NO: 61, respectively.

Examples of other beta-glucosidases useful in the present invention include a *Aspergillus oryzae* beta-glucosidase variant fusion protein of SEQ ID NO: 64 or the *Aspergillus oryzae* beta-glucosidase fusion protein of SEQ ID NO: 66. The beta-glucosidase fusion proteins of SEQ ID NO: 64 and SEQ ID NO: 66 are encoded by SEQ ID NO: 63 and SEQ ID NO: 65, respectively.

The *Aspergillus oryzae* beta-glucosidase can be obtained according to WO 2002/095014. The *Aspergillus fumigatus* beta-glucosidase can be obtained according to WO 2005/047499. The *Penicillium brasilianum* beta-glucosidase can be obtained according to WO 2007/019442. The *Aspergillus niger* beta-glucosidase can be obtained according to Dan *et al.*, 2000, *J. Biol. Chem.* 275: 4973-4980. The *Aspergillus aculeatus* beta-glucosidase can be obtained according to Kawaguchi *et al.*, 1996, *Gene* 173: 287-288.

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.

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.

In one aspect, the 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.

In another aspect, the 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).

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-butyn-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.

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.

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, quinoliny, benzotriazolyl, benzothiazolyl, benzooxazolyl, benzimidazolyl, isoquinoliny, isoindolyl, acridiny, benzoisazolyl, dimethylhydantoin, pyrazinyl, tetrahydrofuranyl, pyrroliny, pyrrolidinyl, morpholinyl, indolyl, diazepiny, azepiny, thiepinyl, piperidinyl, and oxepiny. 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.

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.

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.

The sulfur-containing compound may be any suitable compound comprising one or more sulfur atoms. In one aspect, the sulfur-containing 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.

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.

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.

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, 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.

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), 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).

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; xyl 3 SEQ ID NO: 67 [DNA sequence] and SEQ ID NO: 68 [deduced amino acid sequence]), *Penicillium pinophilum* (WO 2011/041405), *Penicillium* sp. (WO 2010/126772), *Thielavia terrestris* NRRL 8126 (WO 2009/079210), and *Trichophaea saccata* GH10 (WO 2011/057083).

Examples of beta-xylosidases useful in the methods of the present invention include, but are not limited to, *Trichoderma reesei* beta-xylosidase (UniProtKB/TrEMBL accession number Q92458; SEQ ID NO: 69 [DNA sequence] and SEQ ID NO: 70 [deduced amino acid sequence]), *Talaromyces emersonii* (SwissProt accession number Q8X212), and *Neurospora crassa* (SwissProt accession number Q7SOW4).

Examples of acetylxylan esterases useful in the methods of the present invention include, but are not limited to, acetylxylan 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 Q0UHJ1), and *Thielavia terrestris* NRRL 8126 (WO 2009/042846).

5 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).

10 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).

15 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).

20 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).

30 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
35 may be recovered from the fermentation medium and purified by conventional procedures.

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.

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.

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.

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).

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

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*.

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.

5 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*).

10 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. entomophilia*, *C. brassicae*, *C. pseudotropicalis*, *C. boidinii*, *C. utilis*, and *C. scephatae*; *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.*
15 *thermotolerans*, and *K. fragilis*; *Schizosaccharomyces*, such as *S. pombe*; *Thermoanaerobacter*, such as *Thermoanaerobacter saccharolyticum*; and *Zymomonas*, such as *Zymomonas mobilis*.

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
20 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 entomophillia*. In another more preferred aspect, the yeast is *Candida*
25 *pseudotropicalis*. In another more preferred aspect, the yeast is *Candida scephatae*. 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
30 *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
35 yeast is a *Saccharomyces* spp. In a 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*.

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*.

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).

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.

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).

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*.

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

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.

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), which is hereby incorporated by reference.

For ethanol production, following the fermentation the fermented slurry is distilled to extract the ethanol. The ethanol obtained according to the methods of the invention can be used as, e.g., fuel ethanol, drinking ethanol, i.e., potable neutral spirits, or industrial ethanol.

A fermentation stimulator can be used in combination with any of the methods 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), which is hereby incorporated by reference. Examples of

minerals include minerals and mineral salts that can supply nutrients comprising P, K, Mg, S, Ca, Fe, Zn, Mn, and Cu.

Fermentation products: A fermentation product can be any substance derived from the fermentation. The fermentation product can be, without limitation, an alcohol (e.g., 5 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 10 acid, glycine, lysine, serine, and threonine); a gas (e.g., methane, hydrogen (H₂), carbon dioxide (CO₂), and carbon monoxide (CO)); isoprene; a ketone (e.g., acetone); an organic acid (e.g., acetic acid, acetic 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, 15 oxaloacetic acid, propionic acid, succinic acid, and xylonic acid); and polyketide. The fermentation product can also be protein as a high value product.

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 20 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 25 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., 30 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): 35 595-603.

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.

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.

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.

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.

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.

In another preferred aspect, the fermentation product is isoprene.

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*.

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 acetic 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.

In another preferred aspect, the fermentation product is polyketide.

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.

Detergent Compositions

The variant polypeptides having cellulolytic enhancing activity of the present invention may be added to and thus become a component of a detergent composition.

The detergent composition of the present invention may be formulated, for example, as a hand or machine laundry detergent composition including a laundry additive composition suitable for pre-treatment of stained fabrics and a rinse added fabric softener composition, or be formulated as a detergent composition for use in general household hard surface cleaning operations, or be formulated for hand or machine dishwashing operations.

In a specific aspect, the present invention provides a detergent additive comprising a variant polypeptide of the invention. The detergent additive as well as the detergent composition may comprise one or more (*e.g.*, several) enzymes such as a protease, lipase, cutinase, an amylase, carbohydrase, cellulase, pectinase, mannanase, arabinase, galactanase,

xylanase, oxidase, e.g., a laccase, and/or peroxidase.

In general the properties of the selected enzyme(s) should be compatible with the selected detergent, (i.e., pH-optimum, compatibility with other enzymatic and non-enzymatic ingredients, etc.), and the enzyme(s) should be present in effective amounts.

5 Cellulases: Suitable cellulases include those of bacterial or fungal origin. Chemically modified or protein engineered mutants are included. Suitable cellulases include cellulases from the genera *Bacillus*, *Pseudomonas*, *Humicola*, *Fusarium*, *Thielavia*, *Acremonium*, e.g., the fungal cellulases produced from *Humicola insolens*, *Myceliophthora thermophila* and *Fusarium oxysporum* disclosed in US 4,435,307, US 5,648,263, US 5,691,178, US
10 5,776,757 and WO 89/09259.

Especially suitable cellulases are the alkaline or neutral cellulases having color care benefits. Examples of such cellulases are cellulases described in EP 0 495 257, EP 0 531 372, WO 96/11262, WO 96/29397, WO 98/08940. Other examples are cellulase variants such as those described in WO 94/07998, EP 0 531 315, US 5,457,046, US 5,686,593, US
15 5,763,254, WO 95/24471, WO 98/12307 and PCT/DK98/00299.

Commercially available cellulases include CELLUZYME™, and CAREZYME™ (Novozymes A/S), CLAZINASE™, and PURADAX HA™ (Genencor International Inc.), and KAC-500(B)™ (Kao Corporation).

20 Proteases: Suitable proteases include those of animal, vegetable or microbial origin. Microbial origin is preferred. Chemically modified or protein engineered mutants are included. The protease may be a serine protease or a metalloprotease, preferably an alkaline microbial protease or a trypsin-like protease. Examples of alkaline proteases are subtilisins, especially those derived from *Bacillus*, e.g., subtilisin Novo, subtilisin Carlsberg, subtilisin 309, subtilisin 147 and subtilisin 168 (described in WO 89/06279). Examples of
25 trypsin-like proteases are trypsin (e.g., of porcine or bovine origin) and the *Fusarium* protease described in WO 89/06270 and WO 94/25583.

Examples of useful proteases are the variants described in WO 92/19729, WO 98/20115, WO 98/20116, and WO 98/34946, especially the variants with substitutions in one or more of the following positions: 27, 36, 57, 76, 87, 97, 101, 104, 120, 123, 167, 170, 194,
30 206, 218, 222, 224, 235, and 274.

Preferred commercially available protease enzymes include ALCALASE™, SAVINASE™, PRIMASE™, DURALASE™, ESPERASE™, and KANNASE™ (Novozymes A/S), MAXATASE™, MAXACAL™, MAXAPEM™, PROPERASE™, PURAFECT™, PURAFECT OXP™, FN2™, and FN3™ (Genencor International Inc.).

35 Lipases: Suitable lipases include those of bacterial or fungal origin. Chemically modified or protein engineered mutants are included. Examples of useful lipases include

lipases from *Humicola* (synonym *Thermomyces*), e.g., from *H. lanuginosa* (*T. lanuginosus*) as described in EP 258 068 and EP 305 216 or from *H. insolens* as described in WO 96/13580, a *Pseudomonas* lipase, e.g., from *P. alcaligenes* or *P. pseudoalcaligenes* (EP 218 272), *P. cepacia* (EP 331 376), *P. stutzeri* (GB 1,372,034), *P. fluorescens*, *Pseudomonas* sp. strain SD 705 (WO 95/06720 and WO 96/27002), *P. wisconsinensis* (WO 96/12012), a *Bacillus* lipase, e.g., from *B. subtilis* (Dartois *et al.*, 1993, *Biochemica et Biophysica Acta*, 1131: 253-360), *B. stearotherophilus* (JP 64/744992) or *B. pumilus* (WO 91/16422).

Other examples are lipase variants such as those described in WO 92/05249, WO 94/01541, EP 407 225, EP 260 105, WO 95/35381, WO 96/00292, WO 95/30744, WO 94/25578, WO 95/14783, WO 95/22615, WO 97/04079 and WO 97/07202.

Preferred commercially available lipase enzymes include LIPOLASE™ and LIPOLASE ULTRA™ (Novozymes A/S).

Amylases: Suitable amylases (α and/or β) include those of bacterial or fungal origin. Chemically modified or protein engineered mutants are included. Amylases include, for example, α -amylases obtained from *Bacillus*, e.g., a special strain of *Bacillus licheniformis*, described in more detail in GB 1,296,839.

Examples of useful amylases are the variants described in WO 94/02597, WO 94/18314, WO 96/23873, and WO 97/43424, especially the variants with substitutions in one or more of the following positions: 15, 23, 105, 106, 124, 128, 133, 154, 156, 181, 188, 190, 197, 202, 208, 209, 243, 264, 304, 305, 391, 408, and 444.

Commercially available amylases are DURAMYL™, TERMAMYL™, FUNGAMYL™ and BAN™ (Novozymes A/S), RAPIDASE™ and PURASTAR™ (from Genencor International Inc.).

Peroxidases/Oxidases: Suitable peroxidases/oxidases include those of plant, bacterial or fungal origin. Chemically modified or protein engineered mutants are included. Examples of useful peroxidases include peroxidases from *Coprinus*, e.g., from *C. cinereus*, and variants thereof as those described in WO 93/24618, WO 95/10602, and WO 98/15257.

Commercially available peroxidases include GUARDZYME™ (Novozymes A/S).

The detergent enzyme(s) may be included in a detergent composition by adding separate additives containing one or more (e.g., several) enzymes, or by adding a combined additive comprising all of these enzymes. A detergent additive of the invention, i.e., a separate additive or a combined additive, can be formulated, for example, as a granulate, liquid, slurry, etc. Preferred detergent additive formulations are granulates, in particular non-dusting granulates, liquids, in particular stabilized liquids, or slurries.

Non-dusting granulates may be produced, e.g., as disclosed in US 4,106,991 and 4,661,452 and may optionally be coated by methods known in the art. Examples of waxy

coating materials are poly(ethylene oxide) products (polyethyleneglycol, PEG) with mean molar weights of 1000 to 20000; ethoxylated nonylphenols having from 16 to 50 ethylene oxide units; ethoxylated fatty alcohols in which the alcohol contains from 12 to 20 carbon atoms and in which there are 15 to 80 ethylene oxide units; fatty alcohols; fatty acids; and mono- and di- and triglycerides of fatty acids. Examples of film-forming coating materials suitable for application by fluid bed techniques are given in GB 1483591. Liquid enzyme preparations may, for instance, be stabilized by adding a polyol such as propylene glycol, a sugar or sugar alcohol, lactic acid or boric acid according to established methods. Protected enzymes may be prepared according to the method disclosed in EP 238,216.

The detergent composition of the invention may be in any convenient form, e.g., a bar, a tablet, a powder, a granule, a paste or a liquid. A liquid detergent may be aqueous, typically containing up to 70% water and 0-30% organic solvent, or non-aqueous.

The detergent composition comprises one or more (e.g., several) surfactants, which may be non-ionic including semi-polar and/or anionic and/or cationic and/or zwitterionic. The surfactants are typically present at a level of from 0.1% to 60% by weight.

When included therein the detergent will usually contain from about 1% to about 40% of an anionic surfactant such as linear alkylbenzenesulfonate, alpha-olefinsulfonate, alkyl sulfate (fatty alcohol sulfate), alcohol ethoxysulfate, secondary alkanesulfonate, alpha-sulfo fatty acid methyl ester, alkyl- or alkenylsuccinic acid, or soap.

When included therein the detergent will usually contain from about 0.2% to about 40% of a non-ionic surfactant such as alcohol ethoxylate, nonylphenol ethoxylate, alkylpolyglycoside, alkyldimethylamineoxide, ethoxylated fatty acid monoethanolamide, fatty acid monoethanolamide, polyhydroxy alkyl fatty acid amide, or N-acyl N-alkyl derivatives of glucosamine ("glucamides").

The detergent may contain 0-65% of a detergent builder or complexing agent such as zeolite, diphosphate, triphosphate, phosphonate, carbonate, citrate, nitrilotriacetic acid, ethylenediaminetetraacetic acid, diethylenetriaminepentaacetic acid, alkyl- or alkenylsuccinic acid, soluble silicates, or layered silicates (e.g., SKS-6 from Hoechst).

The detergent may comprise one or more (e.g., several) polymers. Examples are carboxymethylcellulose, poly(vinylpyrrolidone), poly(ethylene glycol), poly(vinyl alcohol), poly(vinylpyridine-N-oxide), poly(vinylimidazole), polycarboxylates such as polyacrylates, maleic/acrylic acid copolymers, and lauryl methacrylate/acrylic acid copolymers.

The detergent may contain a bleaching system which may comprise a H₂O₂ source such as perborate or percarbonate which may be combined with a peracid-forming bleach activator such as tetraacetylenediamine or nonanoyloxybenzenesulfonate. Alternatively, the bleaching system may comprise peroxyacids of, for example, the amide, imide, or sulfone type.

The enzyme(s) of the detergent composition of the invention may be stabilized using conventional stabilizing agents, e.g., a polyol such as propylene glycol or glycerol, a sugar or sugar alcohol, lactic acid, boric acid, or a boric acid derivative, e.g., an aromatic borate ester, or a phenyl boronic acid derivative such as 4-formylphenyl boronic acid, and the composition may be formulated as described in, for example, WO 92/19709 and WO 92/19708.

The detergent may also contain other conventional detergent ingredients such as, e.g., fabric conditioners including clays, foam boosters, suds suppressors, anti-corrosion agents, soil-suspending agents, anti-soil redeposition agents, dyes, bactericides, optical brighteners, hydrotropes, tarnish inhibitors, or perfumes.

In the detergent compositions, any enzyme may be added in an amount corresponding to 0.01-100 mg of enzyme protein per liter of wash liquor, preferably 0.05-5 mg of enzyme protein per liter of wash liquor, in particular 0.1-1 mg of enzyme protein per liter of wash liquor.

In the detergent compositions, a variant polypeptide of the present invention having cellulolytic enhancing activity may be added in an amount corresponding to 0.001-100 mg of protein, preferably 0.005-50 mg of protein, more preferably 0.01-25 mg of protein, even more preferably 0.05-10 mg of protein, most preferably 0.05-5 mg of protein, and even most preferably 0.01-1 mg of protein per liter of wash liquor.

A variant polypeptide of the present invention having cellulolytic enhancing activity may also be incorporated in the detergent formulations disclosed in WO 97/07202, which is hereby incorporated by reference.

Plants

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

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

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

Examples of plant parts are stem, callus, leaves, root, fruits, seeds, and tubers as well as the individual tissues comprising these parts, e.g., epidermis, mesophyll,

parenchyme, vascular tissues, meristems. Specific plant cell compartments, such as chloroplasts, apoplasts, mitochondria, vacuoles, peroxisomes and cytoplasm are also considered to be a plant part. Furthermore, any plant cell, whatever the tissue origin, is considered to be a plant part. Likewise, plant parts such as specific tissues and cells isolated to facilitate the utilization of the invention are also considered plant parts, e.g., embryos, endosperms, aleurone and seed coats.

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

The transgenic plant or plant cell expressing a variant 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 (e.g., several) expression constructs encoding a variant into the plant host genome or chloroplast genome and propagating the resulting modified plant or plant cell into a transgenic plant or plant cell.

The expression construct is conveniently a nucleic acid construct that comprises a polynucleotide encoding a variant 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).

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 variant is desired to be expressed. For instance, the expression of the gene encoding a variant 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 Physiol.* 86: 506.

For constitutive expression, the 35S-CaMV, the maize ubiquitin 1, and the rice actin 1 promoter may be used (Franck *et al.*, 1980, *Cell* 21: 285-294; Christensen *et al.*, 1992, *Plant Mol. Biol.* 18: 675-689; Zhang *et al.*, 1991, *Plant Cell* 3: 1155-1165). Organ-specific promoters may be, for example, a promoter from storage sink tissues such as seeds, potato tubers, and fruits (Edwards and Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303), or from metabolic sink tissues such as meristems (Ito *et al.*, 1994, *Plant Mol. Biol.* 24: 863-878), a seed specific promoter such as the glutelin, prolamin, globulin, or albumin promoter from rice (Wu *et al.*, 1998, *Plant Cell Physiol.* 39: 885-889), a *Vicia faba* promoter from the legumin B4 and the unknown seed protein gene from *Vicia faba* (Conrad *et al.*, 1998, *J. Plant Physiol.* 152: 708-711), a promoter from a seed oil body protein (Chen *et al.*, 1998, *Plant Cell*

Physiol. 39: 935-941), the storage protein *napA* promoter from *Brassica napus*, or any other seed specific promoter known in the art, e.g., as described in WO 91/14772. Furthermore, the promoter may be a leaf specific promoter such as the *rbcS* promoter from rice or tomato (Kyoizuka *et al.*, 1993, *Plant Physiol.* 102: 991-1000), the chlorella virus adenine methyltransferase gene promoter (Mitra and Higgins, 1994, *Plant Mol. Biol.* 26: 85-93), the *aldP* gene promoter from rice (Kagaya *et al.*, 1995, *Mol. Gen. Genet.* 248: 668-674), or a wound inducible promoter such as the potato *pin2* promoter (Xu *et al.*, 1993, *Plant Mol. Biol.* 22: 573-588). Likewise, the promoter may be inducible by abiotic treatments such as temperature, drought, or alterations in salinity or induced by exogenously applied substances that activate the promoter, e.g., ethanol, oestrogens, plant hormones such as ethylene, abscisic acid, and gibberellic acid, and heavy metals.

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

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

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

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

Following transformation, the transformants having incorporated the expression construct are selected and regenerated into whole plants according to methods well known

in the art. Often 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.

5 In addition to direct transformation of a particular plant genotype with a construct prepared according to the present invention, transgenic plants may be made by crossing a plant having the construct to a second plant lacking the construct. For example, a construct encoding a variant 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
10 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, or a portion of a DNA construct
15 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 further articulated in U.S. Patent No. 7,151,204.

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

20 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
25 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.

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

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

Examples

Media

2X YT plates were composed of 10 g of tryptone, 5 g of yeast extract, 5 g of sodium chloride, 15 g of Bacto Agar, and deionized water to 1 liter.

PDA plates were composed of 39 g of potato dextrose agar and deionized water to 1 liter.

MDU2BP medium was composed of 45 g of maltose, 1 g of $MgSO_4 \cdot 7H_2O$, 1 g of NaCl, 2 g of K_2HSO_4 , 12 g of KH_2PO_4 , 2 g of urea, 500 μ l of AMG trace metals solution, and deionized water to 1 liter; the pH was adjusted to 5.0 and then filter sterilized with a 0.22 μ m filtering unit.

AMG trace metals solution was composed of 14.3 g of $ZnSO_4 \cdot 7H_2O$, 2.5 g of $CuSO_4 \cdot 5H_2O$, 0.5 g of $NiCl_2 \cdot 6H_2O$, 13.8 g of $FeSO_4 \cdot H_2O$, 8.5 g of $MnSO_4 \cdot 7H_2O$, 3 g of citric acid, and deionized water to 1 liter.

M410 medium was composed of 50 g of maltose, 50 g of glucose, 2 g of $MgSO_4 \cdot 7H_2O$, 2 g of KH_2PO_4 , 4 g of citric acid anhydrous powder, 8 g of yeast extract, 2 g of urea, 0.5 g of AMG trace metals solution, 0.5 g of $CaCl_2$, and deionized water to 1 liter (pH 6.0).

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.

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.

YPG medium was composed of 4 g of yeast extract, 1 g of K_2HPO_4 , 0.5 g of $MgSO_4$, 15.0 g of glucose, and deionized water to 1 liter (pH 6.0).

YPM medium was composed of 1% yeast extract, 2% peptone, and 2% maltodextrin.

SC agar plates were composed of 20 g of agar per liter of SC-URA medium.

SC-URA medium with galactose was composed of 100 ml of 10X Basal salts, 25 ml of 20% casamino acids without vitamins, 10 ml of 1% tryptophan, 4 ml of 5% threonine (filter sterilized, added after autoclaving), and 100 ml of 20% glucose or 100 ml of 20% galactose (filter sterilized, added after autoclaving), and deionized water to 1 liter.

10X Basal salts solution was composed of 75 g of yeast nitrogen base, 113 g of succinic acid, 68 g of NaOH, and deionized water to 1 liter.

YP medium was composed of 10 g of yeast extract, 20 g of Bacto peptone, and deionized water to 1 liter.

COVE salt solution was composed of 26 g of KCl, 26 g of $MgSO_4 \cdot 7H_2O$, 76 g of KH_2PO_4 , 50 ml of COVE trace metals solution, and deionized water to 1 liter.

COVE trace metals solution was composed of 0.04 g of $NaB_4O_7 \cdot 10H_2O$, 0.4 g of

CuSO₄·5H₂O, 1.2 g of FeSO₄·7H₂O, 0.7 g of MnSO₄·H₂O, 0.8 g of Na₂MoO₂·2H₂O, 10 g of ZnSO₄·7H₂O, and deionized water to 1 liter.

COVE plates were composed of 342.3 g of sucrose, 20 ml of COVE salt solution, 10 ml of 1 M acetamide, 10 ml of 1.5 M CsCl, 25 g of Noble agar (Difco), and deionized water to 1 liter.

COVE2 plates were composed of 30 g of sucrose, 20 ml of COVE salt solution, 10 ml of 1 M acetamide, 25 g of Noble agar (Difco), and deionized water to 1 liter.

Trichoderma trace metals solution was composed of 216 g of FeCl₃·6H₂O, 58 g of ZnSO₄·7H₂O, 27 g of MnSO₄·H₂O, 10 g of CuSO₄·5H₂O, 2.4 g of H₃BO₃, 336 g of citric acid, and deionized water to 1 liter.

CIM medium was composed of 20 g of cellulose, 10 g of corn steep solids, 1.45 g of (NH₄)₂SO₄, 2.08 g of KH₂PO₄, 0.28 g of CaCl₂, 0.42 g of MgSO₄·7H₂O, 0.42 ml of *Trichoderma* trace metals solution, 1-2 drops of antifoam, and deionized water to 1 liter; pH adjusted to 6.0.

Example 1: Preparation of *Aspergillus fumigatus* GH61B polypeptide having cellulolytic enhancing activity

A tblastn search (Altschul *et al.*, 1997, *Nucleic Acids Res.* 25: 3389-3402) of the *A. fumigatus* partial genome sequence (The Institute for Genomic Research, Rockville, MD, USA) was performed using as query several known GH61 polypeptides including GH61A polypeptide from *Thermoascus aurantiacus* (GeneSeqP Accession Number AEC05922). Several genes were identified as putative Family GH61 homologs based upon a high degree of similarity to the query sequences at the amino acid level. One genomic region of approximately 850 bp with greater than 70% sequence identity to the *Thermoascus aurantiacus* GH61A polypeptide sequence at the amino acid level was chosen for further study.

A. fumigatus NN051616 was grown and harvested as described in U.S. Patent No. 7,244,605. Frozen mycelia were ground, by mortar and pestle, to a fine powder and genomic DNA was isolated using a DNEASY® Plant Maxi Kit (QIAGEN Inc., Valencia, CA, USA) according to manufacturer's instructions.

Two synthetic oligonucleotide primers shown below were designed to PCR amplify the *A. fumigatus* Family GH61B polypeptide gene from the genomic DNA. An IN-FUSION® Cloning Kit (BD Biosciences, Palo Alto, CA, USA) was used to clone the fragment directly into the expression vector pAILo2 (WO 2004/099228), without the need for restriction digestion and ligation.

Forward primer:

5'-ACTGGATTTACCATGACTTTGTCCAAGATCACTTCCA-3' (SEQ ID NO: 71)

Reverse primer:

5'-TCACCTCTAGTTAATTAAGCGTTGAACAGTGCAGGACCAG-3' (SEQ ID NO: 72)

Bold letters represent coding sequence. The remaining sequence is homologous to the insertion sites of pALo2.

5 Fifty picomoles of each of the primers above were used in a PCR reaction composed of 204 ng of *A. fumigatus* genomic DNA, 1X *Pfx* Amplification Buffer (Invitrogen Corp., Carlsbad, CA, USA), 1.5 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 2.5 units of PLATINUM® *Pfx* DNA polymerase (Invitrogen Corp., Carlsbad, CA, USA), and 1 µl of 50 mM MgSO₄ in a final volume of 50 µl. The amplification was performed using an
10 EPPENDORF® MASTERCYCLER® 5333 eppgradient S (Eppendorf Scientific, Inc., Westbury, NY, USA) programmed for 1 cycle at 94°C for 3 minutes; and 30 cycles each at 94°C for 30 seconds, 56°C for 30 seconds, and 72°C for 1 minutes. The heat block was then held at 72°C for 15 minutes followed by a 4°C soak cycle.

The reaction products were isolated by 1.0% agarose gel electrophoresis using 40
15 mM Tris base-20 mM sodium acetate-1 mM disodium EDTA (TAE) buffer where an approximately 850 bp product band was excised from the gel and purified using a MINELUTE® Gel Extraction Kit (QIAGEN Inc., Valencia, CA, USA) according to the manufacturer's instructions.

The fragment was then cloned into pALo2 using an IN-FUSION® Cloning Kit. The
20 vector was digested with *Nco* I and *Pac* I. The fragment was purified by gel electrophoresis as above and a QIAQUICK® Gel Purification Kit (QIAGEN Inc., Valencia, CA, USA). The gene fragment and the digested vector were combined together in a reaction resulting in the expression plasmid pAG43 in which transcription of the Family GH61B polypeptide gene was under the control of the NA2-tpi promoter. The NA2-tpi promoter is a modified promoter
25 from the *Aspergillus niger* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from the *Aspergillus nidulans* triose phosphate isomerase gene. The recombination reaction (20 µl) was composed of 1X IN-FUSION® Buffer (BD Biosciences, Palo Alto, CA, USA), 1X BSA (BD Biosciences, Palo Alto, CA, USA), 1 µl of IN-FUSION® enzyme (diluted 1:10) (BD Biosciences, Palo Alto, CA, USA), 166
30 ng of pALo2 digested with *Nco* I and *Pac* I, and 110 ng of the *A. fumigatus* GH61B polypeptide purified PCR product. The reaction was incubated at 37°C for 15 minutes followed by 15 minutes at 50°C. The reaction was diluted with 40 µl of a 10 mM Tris-0.1 M EDTA buffer and 2.5 µl of the diluted reaction were used to transform *E. coli* XL10 SOLOPACK® Gold competent cells (Stratagene, La Jolla, CA, USA). An *E. coli* transformant
35 containing pAG43 (GH61B protein gene) was identified by restriction enzyme digestion and plasmid DNA was prepared using a BIOROBOT® 9600 (QIAGEN Inc., Valencia, CA, USA).

DNA sequencing of the 862 bp PCR fragment was performed with an Applied

Biosystems Model 377 XL Automated DNA Sequencer (Applied Biosystems, Carlsbad, CA, USA) using dye-terminator chemistry (Giesecke *et al.*, 1992, *Journal of Virology Methods* 38: 47-60) and primer walking strategy. The following vector specific primers were used for sequencing:

5 pAllo2 5 Seq:

5'-TGTCCTTGTTCGATGCG 3' (SEQ ID NO: 73)

pAllo2 3 Seq:

5'-CACATGACTTGGCTTCC 3' (SEQ ID NO: 74)

10 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).

A gene model for the *A. fumigatus* sequence was constructed based on similarity of the encoded protein to the *Thermoascus aurantiacus* GH61A polypeptide (GeneSeqP Accession Number AEC05922). The nucleotide sequence (SEQ ID NO: 1) and deduced amino acid sequence (SEQ ID NO: 2) of the *A. fumigatus* GH61B polypeptide gene are shown in Figures 1A and 1B. The genomic fragment encodes a polypeptide of 250 amino acids, interrupted by 2 introns of 53 and 56 bp. The % G+C content of the gene and the mature coding sequence are 53.9% and 57%, respectively. Using the SignalP software program (Nielsen *et al.*, 1997, *Protein Engineering* 10: 1-6), a signal peptide of 21 residues was predicted. The predicted mature protein contains 221 amino acids with a predicted molecular mass of 23.39 kDa.

25 *Aspergillus oryzae* JaL355 protoplasts were prepared according to the method of Christensen *et al.*, 1988, *Bio/Technology* 6: 1419-1422. Six µg of pAG43 were used to transform *Aspergillus oryzae* JaL355. Twenty-six transformants were isolated to individual PDA plates.

Confluent PDA plates of 24 transformants were each washed with 5 ml of 0.01% TWEEN® 20 and the spores were each collected. Eight µl of each spore stock were added to 1 ml of YPG, YPM, and M410 media separately in 24 well plates and incubated at 34°C. After 3 days of incubation, 7.5 µl of supernatant from four transformants were analyzed using a CRITERION® stain-free, 8-16% gradient SDS-PAGE gel (Bio-Rad Laboratories, Inc., Hercules, CA, USA) according to the manufacturer's instructions. Based on this gel, M410 medium was chosen as the best medium. Five days after incubation, 7.5 µl of supernatant from each M410 culture were analyzed using a CRITERION® stain-free, 8-16% gradient SDS-PAGE gel. SDS-PAGE profiles of the cultures showed that several transformants had a new major band of approximately 25 kDa.

A confluent plate of one transformant (grown on a PDA plate) was washed with 5 ml of 0.01% TWEEN® 20 and inoculated into four 500 ml Erlenmeyer flasks containing 100 ml

of M410 medium to generate broth for characterization of the enzyme. The flasks were harvested on day 5 (300 ml), filtered using a 0.22 µm EXPRESS™ Plus Membrane (Millipore, Bedford, MA, USA), and stored at 4°C.

5 The filtered shake flask broth containing the recombinantly produced *A. fumigatus* GH61B polypeptide having cellulolytic enhancing activity was first concentrated by a tangential flow concentrator (Pall Filtron, Northborough, MA, USA) equipped with a 10 kDa polyethersulfone membrane (Pall Filtron, Northborough, MA, USA), buffer exchanged into 20 mM Tris-HCl pH 8.0, and then purified using a HIGHLOAD™ 26/60 SUPERDEX™ 75 gel filtration column (GE Healthcare, Piscataway, NJ, USA) with a 750 ml isocratic gradient in
10 150 mM NaCl, 20 mM Tris-HCl pH 8.0. Fractions were collected and pooled based on SDS-PAGE analysis. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit (Thermo Fisher Scientific Inc., Rockford, IL, USA) in which bovine serum albumin was used as a protein standard.

15 **Example 2: Construction of *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V**

A variant of the wild-type *Aspergillus fumigatus* GH61B polypeptide having cellulolytic enhancing activity was constructed with the substitutions I75V, F77L, F179I, I181L, and I183V. The wild-type *A. fumigatus* GH61B backbone from plasmid pAG43 (Example 2) was
20 used as starting template upon which substitutions were introduced in several separate steps, generating intermediates until the following plasmid with the final target amino acid substitutions was obtained: pTH227 (I75V + F77L + F179I + I181L + I183V). A QUIKCHANGE® XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) was used to generate the mutations via PCR mediated reactions in which synthetic
25 oligonucleotide primer pairs, as shown in Table 1, designed with the target changes were used to incorporate the desired substitutions.

Variant plasmid pTH221 was generated using the reverse and forward primer pair 68337 and 68343 and pAG43 as the starting template. Variant plasmid pTH227 was generated using the reverse and forward primer pair 68335 and 68336 and variant plasmid
30 pTH221 as the starting template.

Table 1

<u>Amino acid changes</u>	<u>Primer ID</u>	<u>Sequences</u>	<u>Plasmid name</u>
F179I, I181L, I183V	68337	GCAGAACTACCCCCAGTGTATCAACCTGCAAGTGACCGGT GGCGGCAGTGCTCAGG (SEQ ID NO: 75)	pTH221
	68343	CCTGAGCACTGCCGCCACCGGTCACTTGCAGGTTGATACA CTGGGGGTAGTTCTGC (SEQ ID NO: 76)	
I75V, F77L	68335	GCAGCCGGTTCACAGGTGGAAGTGCAGTGGACGACGTGG (SEQ ID NO: 77)	pTH227
	68336	CCACGTCGTCCACTGCAGTTCACCTGTGAACCGGCTGC (SEQ ID NO: 78)	

The resulting mutant plasmid DNAs were prepared using a BIOROBOT® 9600 and sequenced using a 3130xl Genetic Analyzer (Applied Biosystems, Carlsbad, CA, USA).

5

Example 3: Expression of the *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V in *Aspergillus oryzae* JaL250

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 pTH227 (or pAllo2 as a control). The transformation yielded about 20-25 transformants. The transformants were then spore purified to individual selective PDA plates and then grown in 24 well culture plates containing 1 ml of MDU2BP medium and incubated at 34°C stationary for 5 days. Broth samples were harvested at day 5 and analyzed by 8-16% Tris-glycine SDS-PAGE (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Once the cultures from each spore purified transformant were confluent and had sporulated, spore stocks were made by applying 5 ml of sterile filtered 0.01% TWEEN® 80 (diluted with glass distilled water) onto the center of each PDA plate and using a sterile spreader to scrape the spores into solution. Spore stocks from the highest producing transformants for each batch identified by SDS-PAGE as having darker bands at the predicted molecular weight of 26 kDa were used to inoculate a 2 liter shake flask containing 300 ml of MDU2BP medium. Shake flasks were incubated for 5 days at 34°C with agitation at 220 rpm. After the incubation, the broths were sterile filtered using a 0.22 µm polyethersulfone membrane (Millipore, Bedford, MA, USA) for purification. The *A. oryzae* strain identified from SDS-PAGE analysis of the shake flask broths with the strongest band at 26 kDa was TH168 (*Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V).

Example 4: Purification of the *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V

The recombinantly produced *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V was first concentrated by a tangential flow concentrator (Pall Filtron, Northborough, MA, USA) equipped with a 10 kDa polyethersulfone membrane (Pall Filtron, Northborough, MA, USA), buffer exchanged into 20 mM Tris-HCl pH 8.0, and then purified using a (self-packed) 75 ml Q-SEPHAROSE® High Performance column (GE Healthcare, Piscataway, NJ, USA) with a 800 ml 0-600 mM NaCl linear gradient in 20 mM Tris-HCl pH 8.0. Fractions were collected and pooled based on SDS-PAGE analysis. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit in which bovine serum albumin was used as a protein standard.

Example 5: Preparation of *Trichoderma reesei* CEL5A endoglucanase II

The *Trichoderma reesei* Family GH5A endoglucanase II gene (SEQ ID NO: 5 [DNA sequence] and SEQ ID NO: 6 [deduced amino acid sequence]) was cloned into an *Aspergillus oryzae* expression vector as described below.

Two synthetic oligonucleotide primers, shown below, were designed to PCR amplify the endoglucanase II gene from *T. reesei* RutC30 genomic DNA. Genomic DNA was isolated using a DNEASY® Plant Maxi Kit. An IN-FUSION™ PCR Cloning Kit was used to clone the fragment directly into pAILo2 (WO 2004/099228).

Forward primer:

5'-ACTGGATTTACCAT**GAACAAGTCCGTGGCTCCATTGCT**-3' (SEQ ID NO: 79)

Reverse primer:

5'-TCACCTCTAGTTAATTA**ACTACTTTCTTGCGAGACACG**-3' (SEQ ID NO: 80)

Bold letters represent coding sequence. The remaining sequence contains sequence identity compared with the insertion sites of pAILo2 (WO 2004/099228).

Fifty picomoles of each of the primers above were used in a PCR reaction composed of 200 ng of *T. reesei* genomic DNA, 1X Pfx Amplification Buffer, 6 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 2.5 units of PLATINUM® Pfx DNA polymerase, and 1 µl of 50 mM MgSO₄ in a final volume of 50 µl. The amplification reaction was incubated in an EPPENDORF® MASTERCYCLER® 5333 (Eppendorf Scientific, Inc., Westbury, NY, USA) 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 at 10°C. A 1.5 kb PCR reaction product was isolated on a 0.8% GTG® agarose gel (Cambrex Bioproducts, 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™ (Clare Chemical Research, Dolores, CO, USA). The 1.5 kb DNA band was excised with a disposable razor blade and purified using an ULTRAFREE® DA spin cup (Millipore, Billerica, MA, USA) according to the manufacturer's instructions.

Plasmid pAILo2 was linearized by digestion with *Nco* I and *Pac* I. The plasmid 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 using an IN-FUSION™ PCR Cloning Kit. The reaction (20 µl) contained 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 100 ng of the *T. reesei* CEL5A endoglucanase II 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* XL10 SOLOPACK® Gold competent cells according to the manufacturer's instructions. After a recovery period, two 100 µl aliquots from the transformation reaction were plated onto 150 mm 2X YT plates supplemented with 100 µg of ampicillin per ml. The plates were incubated overnight at 37°C. A set of 3 putative recombinant clones was recovered from the selection plates and plasmid DNA was prepared from each one using a BIOROBOT® 9600. Clones were analyzed by *Pci* I/*Bsp* LU11 I restriction digestion. One clone with the expected restriction digestion pattern was then sequenced to confirm that there were no mutations in the cloned insert. Clone #3 was selected and designated pAILo27.

Aspergillus oryzae JaL250 (WO 99/61651) protoplasts were prepared according to the method of Christensen *et al.*, 1988, *supra*, and transformed with 5 µg of pAILo27 (or pAILo2 as a control). The transformation yielded about 50 transformants. Eleven transformants were isolated to individual PDA plates and incubated for five days at 34°C.

Confluent spore plates were washed with 3 ml of 0.01% TWEEN® 80 and the spore suspension was used to inoculate 25 ml of MDU2BP medium in 125 ml glass shake flasks. Transformant cultures were incubated at 34°C with constant shaking at 200 rpm. At day five post-inoculation, cultures were centrifuged at 6000 x *g* and their supernatants collected. Five microliters of each supernatant were mixed with an equal volume of 2X loading buffer (10% beta-mercaptoethanol) and loaded onto a 1.5 mm 8%-16% Tris-Glycine SDS-PAGE gel and stained with SIMPLYBLUE™ SafeStain (Invitrogen Corp., Carlsbad, CA, USA). SDS-PAGE profiles of the culture broths showed that ten out of eleven transformants produced a new protein band of approximately 45 kDa. Transformant number 1, designated *Aspergillus oryzae* JaL250AILo27, was cultivated in a fermentor.

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.

One hundred ml of shake flask medium were added to a 500 ml shake flask. The shake flask was inoculated with two plugs of *A. oryzae* JaL250AILo27 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.

5 Fermentation batch medium was composed 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, 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. Fermentation feed medium was composed of maltose.

10 A total of 1.8 liters of the fermentation batch medium was added to an Applikon Biotechnology three liter glass jacketed fermentor (Applikon Biotechnology, Inc., Foster City, CA, USA). 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, Inc., Foster City, CA, USA) to a set-point of 6.1 +/- 0.1. Air was added to the vessel at a rate of 1 vvm and the 15 broth was agitated by 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 using a 0.22 µm EXPRESS™ Plus Membrane, and stored at 4°C.

20 The supernatant was desalted and buffer-exchanged into 20 mM Tris-HCl pH 8.0 using a 400 ml SEPHADEX™ G25 desalting column (GE Healthcare, Piscataway, NJ, USA) according to the manufacturer's instructions. Desalted *Trichoderma reesei* CEL5A endoglucanase II was loaded onto a MONOQ™ HR 16/10 ion exchange column (GE Healthcare, Piscataway, NJ, USA) and eluted with a 300 ml linear 0-300 mM NaCl gradient in 20 mM Tris-HCl pH 8 with collection of 10 ml fractions. Fractions were pooled based on 25 SDS-PAGE analysis. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit in which bovine serum albumin was used as a protein standard.

Example 6: Preparation of *Aspergillus fumigatus* GH3A beta-glucosidase

30 An *A. fumigatus* beta-glucosidase (SEQ ID NO: 55 [DNA sequence] and SEQ ID NO: 56 [deduced amino acid sequence]) was prepared according to U.S. Patent No. 7,244,605. 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 *Aspergillus fumigatus* NN055679 Cel7A cellobiohydrolase I

35 A tfasty search (Pearson *et al.*, 1997, *Genomics* 46:24-36) of the *A. fumigatus* partial genome sequence (The Institute for Genomic Research, Rockville, MD, USA) 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 sequence identity to the query sequence was chosen for further study, and the corresponding gene was named *cel7A*.

5 Two synthetic oligonucleotide primers shown below were designed to PCR amplify the *A. fumigatus* NN055679 *cel7A* cellobiohydrolase I gene (SEQ ID NO: 49 [DNA sequence] and SEQ ID NO: 50 [deduced amino acid sequence]) from genomic DNA of *A. fumigatus* prepared as described in WO 2005/047499.

Forward primer:

10 5'-gggcATGCTGGCCTCCACCTTCTCC-3' (SEQ ID NO: 81)

Reverse primer:

5'-gggtaattaaCTACAGGCACTGAGAGTAA-3' (SEQ ID NO: 82)

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 *A. 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.

The fragment was digested with *Sph* I and *Pac* I and ligated into the expression vector pAILo2 also digested with *Sph* I and *Pac* I according to standard procedures. The ligation products were transformed into *E. coli* XL10 SOLOPACK® Gold competent cells 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. DNA sequencing of the insert from this plasmid was performed with a Perkin-Elmer 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, *supra*) 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 (SEQ ID NO: 49 [DNA sequence] and SEQ ID NO: 50 [deduced amino acid sequence]). The resulting plasmid was named pEJG93.

Aspergillus oryzae JaL250 protoplasts were prepared according to the method of Christensen *et al.*, 1988, *supra*, and transformed with 5 µg of pEJG93 (as well as pAILo2 as a vector control). The transformation yielded about 100 transformants. Ten transformants were isolated to individual PDA plates.

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 Corp., 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 *A. oryzae* JaL250EJG93.

One hundred ml of shake flask medium (Example 5) were added to a 500 ml shake flask. The shake flask was inoculated with two plugs of *A. 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 was used to inoculate a 3 liter fermentation vessel.

A total of 1.8 liters of the fermentation batch medium (Example 5) was added to an Applikon Biotechnology three liter glass jacketed fermentor. Fermentation feed medium (Example 5) 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 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 using a 0.22 µm EXPRESS™ Plus Membrane, and stored at 4°C.

Filtered broth was concentrated and buffer exchanged using a tangential flow concentrator (Pall Filtron, Northborough, MA, USA) equipped with a 10 kDa polyethersulfone membrane (Pall Filtron, Northborough, MA, USA) with 20 mM Tris-HCl pH 8., and then purified using a (self-packed) 75 ml Q-SEPHAROSE® High Performance column (with a 750 ml 0-600 mM NaCl linear gradient in 20 mM Tris-HCl pH 8.0 with collection of 10 ml fractions. Fractions pooled based on SDS-PAGE analysis. 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* Cel6A cellobiohydrolase II

A. fumigatus NN055679 cellobiohydrolase II (CBHII) (SEQ ID NO: 51 [DNA sequence] and SEQ ID NO: 52 [deduced amino acid sequence]) was prepared according to the following procedure.

Two synthetic oligonucleotide primers, shown below, were designed to PCR amplify the full-length open reading frame of the *A. fumigatus* Family 6A glycosyl hydrolase 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'-ACTGGATTTACC**ATGAAGCACCTTGCATCTTCCATCG**-3' (SEQ ID NO: 83)

Reverse primer:

5'-TCACCTCTAGTTAATTA**AAAGGACGGGTTAGCGT**-3' (SEQ ID NO: 84)

- 5 Bold letters represent coding sequence. The remaining sequence contains sequence identity compared with the insertion sites of pAIIo2.

Fifty picomoles of each of the primers above were used in a PCR reaction composed of 500 ng of *A. 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, 0.1
10 unit of *Taq* DNA polymerase (New England Biolabs, Ipswich, MA, USA), 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 96°C for 30 seconds, 61°C for 30 seconds, and 72°C for 2 minutes. After the 35 cycles, the reaction was
15 incubated at 72°C for 10 minutes and then cooled at 10°C until further processed. To remove the A-tails produced by *Taq* DNA polymerase the reaction was incubated for 10 minutes at 68°C in the presence of 1 unit of *Pfx* DNA polymerase (Invitrogen Corp., Carlsbad, CA, USA).

A 1.3 kb PCR reaction product was isolated on a 0.8% GTG-agarose gel (Cambrex Bioproducts, East Rutherford, NJ, USA) using TAE buffer and 0.1 µg of ethidium bromide
20 per ml. The DNA band was visualized with the aid of a DARK READER™ (Clare Chemical Research, Dolores, CO, USA) to avoid UV-induced mutations. The 1.3 kb DNA band was excised with a disposable razor blade and purified with an ULTRAFREE® DA spin cup according to the manufacturer's instructions.

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

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 digest.
35 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 clones 2, 4, 5, 6, 7 and 8 were 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 μ l of 1:10 diluted 0.1 mM EDTA-10 mM Tris pH 7.4 and 1 μ l of this mix was used to re-amplify the *A. fumigatus* glycosyl hydrolase 6A coding region.

Fifty picomoles of each of the primers above were used in a PCR reaction composed of 1 μ l of the diluted mix of clones 2, 6 and 7, 1X *Pfx* Amplification Buffer, 6 μ l of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 2.5 units of PLATINUM® *Pfx* DNA polymerase, 1 μ l of 50 mM MgSO₄, in a final volume of 50 μ l. 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 at 10°C until further processed. A 1.3 kb PCR reaction product was isolated on a 0.8% GTG-agarose gel using TAE buffer and 0.1 μ g 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 from the gel with a disposable razor blade and purified with an ULTRAFREE® DA spin cup according to the manufacturer's instructions.

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 μ l) was composed of 1X IN-FUSION™ Buffer, 1X BSA, 1 μ l of IN-FUSION™ enzyme (diluted 1:10), 100 ng of pAILo2 digested with *Nco* I and *Pac* I, and 50 ng of the *A. fumigatus* GH6A purified PCR product. The reaction was incubated at room temperature for 30 minutes. A 2 μ l sample of the reaction was used to transform transform *E. coli* TOP10 competent cells according to the manufacturer's instructions. After the recovery period, two 100 μ l aliquots from the transformation reaction were plated onto 150 mm 2X YT plates supplemented with 100 μ g 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 pAILo33.

Aspergillus oryzae JaL355 protoplasts were prepared according to the method of Christensen *et al.*, 1988, *supra*, and transformed with 5 μ g of pAILo33 (as well as pAILo2 as a vector control). The transformation yielded about 100 transformants. Ten transformants were isolated to individual PDA plates.

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 Corp., 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 *A. oryzae* JaL355 ALlo33.

One hundred ml of shake flask medium (Example 5) was added to a 500 ml shake flask. The shake flask was inoculated with two plugs of *A. oryzae* JaL355 ALlo33 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 was used to inoculate a 3 liter fermentation vessel.

A total of 1.8 liters of the fermentation batch medium (Example 5) was added to an Applikon Biotechnology three liter glass jacketed fermentor. Fermentation feed medium (Example 5) 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 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 using a 0.22 µm EXPRESS™ Plus Membrane, and stored at 4°C.

The broth was filtered using a 0.7 µm glass filter GF/F (Whatman, Piscataway, NJ, USA) and then using a 0.22 µm EXPRESS™ Plus Membrane. The filtered broth was concentrated and buffer exchanged using a tangential flow concentrator (Pall Filtron, Northborough, MA, USA) equipped with a 10 kDa polyethersulfone membrane (Pall Filtron, Northborough, MA, USA) with 20 mM Tris-HCl pH 8.0.

Desalted broth was loaded onto on a MONOQ™ HR 16/10 ion exchange column and eluted with a 300 ml linear 0-300 mM NaCl gradient in 20 mM Tris-HCl pH 8 with collection of 10 ml fractions. Fractions pooled based on SDS-PAGE analysis. Pooled fractions were adjusted to 1.2 M (NH₄)₂SO₄ in 20 mM Tris, pH 8.0 and then applied to a 75 ml self-poured PHENYL SEPHAROSE™ FAST-FLOW® HIGH-SUB® column (GE Healthcare, Piscataway, NJ, USA) and eluted with a 750 ml linear 1.2-0 M (NH₄)₂SO₄ gradient in 20 mM Tris-HCl pH 8 with collection of 10 ml fractions. Fractions pooled based on SDS-PAGE were concentrated using a VIVASPIN® 10 kDa MWCO centrifugal concentrator (GE Healthcare, Piscataway, NJ, USA). Concentrated material was then further purified using a HIGHLOAD™ 26/60 Superdex 75 gel filtration column with a 200 ml isocratic gradient in 150 mM NaCl, 20 mM Tris-HCl pH 8.0. Fractions were collected and pooled based on SDS-PAGE and

concentrated using concentrated using VIVASPIN® 10 kDa MWCO centrifugal concentrators. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit in which bovine serum albumin was used as a protein standard.

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

Aspergillus fumigatus NN055679 GH10 xylanase (xyn3; SEQ ID NO: 67 [DNA sequence] and SEQ ID NO: 68 [deduced amino acid sequence]) was prepared recombinantly according to WO 2006/078256 using *Aspergillus oryzae* BECh2 as a host.

10 The filtered broth was desalted and buffer-exchanged using a tangential flow concentrator (Pall Filtron, Northborough, MA, USA) equipped with a 10 kDa polyethersulfone membrane (Pall Filtron, Northborough, MA, USA) with 25 mM Tris-HCl pH 8.5. Desalted material was applied to 400 ml (hand-poured) Q-SEPHAROSE™ Fast-Flow™ column (GE Healthcare, Piscataway, NJ, USA), washed with 400 ml 25 mM Tris pH 8.5, and eluted with a 1400 ml linear gradient of 0-600 mM NaCl in 25 mM Tris pH 8.5 with collection of 10 ml
15 fractions. Fractions were pooled based on SDS-PAGE, and adjusted to 1.5 M (NH₄)₂SO₄, 20 mM Tris pH 8.0 and applied to a 75 ml self-poured PHENYL SEPHAROSE™ FAST-FLOW® HIGH-SUB® column, washed with 75 ml 20 mM Tris-HCl pH 8.0, and eluted with a 1500 ml linear 1.5-0 M (NH₄)₂SO₄ gradient in 20 mM Tris-HCl pH 8 with collection of 10 ml fractions. Fractions were pooled based on SDS-PAGE analysis and concentrated using a 300 ml
20 stirred cell concentration device (Millipore, Bedford, MA, USA) equipped with a 10 kDa MWCO membrane and desalted into 20 mM Tris-HCl pH 8.0, 150 mM NaCl. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit with bovine serum albumin as a protein standard.

25 **Example 10: Preparation of *Trichoderma reesei* RutC30 GH3 beta-xylosidase**

A *Trichoderma reesei* RutC30 beta-xylosidase gene (SEQ ID NO: 69 [DNA sequence] and SEQ ID NO: 70 [deduced amino acid sequence]) was isolated by screening a Lambda ZAP®-CMR XR Library prepared from *T. reesei* RutC30 genomic DNA using a Lambda ZAP®-CMR XR Library Construction Kit (Stratagene, La Jolla, CA, USA) according
30 to the manufacturer's instructions. *T. reesei* RutC30 genomic DNA was prepared using standard methods. A DNA segment encoding 2300 bp of the *T. reesei* beta-xylosidase was amplified using the PCR primers shown below.

Forward Primer:

5'-gtgaataacgcagctcttctcg-3' (SEQ ID NO: 85)

35 Reverse Primer:

5'-ccttaattaattatgcgtcaggtgt-3' (SEQ ID NO: 86)

Primer 994768 was designed to amplify from the first base after the beta-xylosidase

start site and primer 994769 was designed with a *Pac* I site at the 5' end.

Fifty picomoles of each of the primers above were used in a PCR reaction composed of 50 ng of plasmid DNA from the lamda zap library, 1 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 5 µl of 10X PLATINUM® *Pfx* DNA Polymerase Buffer, and 1 unit of PLATINUM® *Pfx* DNA polymerase, in a final volume of 50 µl. An EPPENDORF® MASTERCYCLER® 5333 was used to amplify the DNA fragment programmed for 1 cycle at 95°C for 3 minutes; and 30 cycles each at 94°C for 45 seconds, 55°C for 60 seconds, and 72°C for 1 minute 30 seconds. After the 30 cycles, the reaction was incubated at 72°C for 10 minutes and then cooled to 4°C until further processing.

A 2.3 kb PCR product was purified by 1% agarose gel electrophoresis using TAE buffer, excised from the gel, and purified using a QIAQUICK® Gel Extraction Kit. The 2.3 kb PCR product was then digested with *Pac* I to facilitate insertion into pAIlO1 (WO 2004/099228).

The pAIlO1 vector was digested with *Nco* I and then filled in using T4 DNA polymerase (Roche Applied Science, Nutley, NJ, USA) according to manufacturer's instructions. A second enzyme, *Pac* I, was then used to digest the 5' end of pAIlO1 and the reaction was purified by agarose gel electrophoresis as described above to isolate a 6.9 kb vector fragment.

The 2.3 kb beta-xylosidase fragment was then ligated to the 6.9 kb vector fragment and transformed into *E. coli* XL1-Blue Subcloning Competent Cells (Invitrogen Corp., Carlsbad, CA, USA) according to manufacturer's instructions. Transformants were screened using restriction digestion analysis in order to identify those with the correct insert. A new expression vector, pSaMe04, was confirmed by sequencing using an ABI 3700 DNA Analyzer (Applied Biosystems, Foster City, CA, USA) and dye terminator chemistry (Giesecke *et al.*, 1992, *supra*).

Two synthetic oligonucleotide primers shown below were designed to PCR amplify the *Trichoderma reesei* beta-xylosidase gene from pSaMe04 to construct a *Trichoderma* expression vector. An IN-FUSION™ Cloning Kit was used to clone the fragment directly into the expression vector pMJ09 (WO 2005/056772), without the need for restriction digestion and ligation.

TrBXYL-F (ID 064491):

5'-CGGACTGCGCACCC**ATGGTGAATAACGCAGCTCT**-3' (SEQ ID NO: 87)

TrBXYL-R (ID 064492):

5'-TCGCCACGGAGCTT**ATTATGCGTCAGGTGTAGCAT**-3' (SEQ ID NO: 88)

Bold letters represent coding sequence. The remaining sequence is homologous to the insertion sites of pMJ09.

Fifty picomoles of each of the primers above were used in a PCR reaction composed

of 50 ng of pSaMe04, 1 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 5 µl of 10X ACCUTAQ™ DNA Polymerase Buffer (Sigma-Aldrich, St. Louis, MO, USA), and 5 units of ACCUTAQ™ DNA polymerase (Sigma-Aldrich, St. Louis, MO, USA), in a final volume of 50 µl. An EPPENDORF® MASTERCYCLER® 5333 was used to amplify the DNA fragment
5 programmed for 1 cycle at 95°C for 3 minutes; and 30 cycles each at 94°C for 45 seconds, 55°C for 60 seconds, and 72°C for 1 minute 30 seconds. After the 30 cycles, the reaction was incubated at 72°C for 10 minutes and then cooled to 4°C until further processing.

The reaction products were isolated by 1.0% agarose gel electrophoresis using TAE buffer where a 1.2 kb product band was excised from the gel and purified using a
10 QIAQUICK® Gel Extraction Kit according to the manufacturer's instructions.

The 1.2 kb fragment was then cloned into pMJ09 using an IN-FUSION™ Cloning Kit. The vector was digested with *Nco* I and *Pac* I and purified by agarose gel electrophoresis as described above. The gene fragment and the digested vector were ligated together in a reaction resulting in the expression plasmid pSaMe-TrBXYL in which transcription of the
15 beta-xylosidase gene was under the control of the *T. reesei cbh1* gene promoter. The ligation reaction (50 µl) was composed of 1X IN-FUSION™ Buffer, 1X BSA, 1 µl of IN-FUSION™ enzyme (diluted 1:10), 100 ng of pMJ09 digested with *Nco* I and *Pac* I, and 100 ng of the *T. reesei* beta-xylosidase purified PCR product. The reaction was incubated at room temperature for 30 minutes. One µl of the reaction was used to transform *E. coli* XL10
20 SOLOPACK® Gold competent cells. An *E. coli* transformant containing pSaMe-TrBXYL was detected by restriction enzyme digestion and plasmid DNA was prepared using a BIOROBOT® 9600. DNA sequencing of the *T. reesei* beta-xylosidase gene from pSaMe-TrBXYL was performed using dye-terminator chemistry (Giesecke *et al.*, 1992, *supra*) and primer walking strategy.

25 Plasmid pSaMe-AaXYL was constructed to comprise the *Trichoderma reesei* cellobiohydrolase I gene promoter and terminator and the *Aspergillus aculeatus* GH10 xylanase coding sequence.

Cloning of the *A. aculeatus* xylanase followed the overall expression cloning protocol as outlined in H. Dalbøge *et al.*, 1994, *Mol. Gen. Genet.* 243: 253-260.

30 RNA was isolated from *A. aculeatus* CBS 101.43 mycelium. Poly(A)⁺ RNA was isolated from total RNA by chromatography on oligo(dT)-cellulose. Double-stranded cDNA was synthesized as described by Maniatis *et al.* (Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, 1982). After synthesis the cDNA was treated with mung bean nuclease, blunt-ended with T4 DNA polymerase, and ligated to non-palindromic
35 *Bst* XI adaptors (Invitrogen Corp., Carlsbad, CA, USA). The cDNA was size fractionated by 1% agarose gel electrophoresis using TAE buffer where fragments ranging from 600 bp to 4000 bp were used in the library construction. The DNA was ligated into *Bst* XI-digested

pYES 2.0 (Invitrogen Corp., Carlsbad, CA, USA) between the GAL1 promoter and the iso-1-cytochrome c terminator and transformed into *Escherichia coli* MC1061 cells (Stratagene, La Jolla, CA, USA). The library was plated onto LB plates and incubated overnight at 37°C. The colonies were scraped from the plates and resuspended in LB medium supplemented with
5 100 µg of ampicillin per ml. Plasmid DNA was isolated using a Plasmid Midi Kit (QIAGEN Inc., Valencia, CA, USA). The purified plasmid DNA was pooled.

The purified plasmid DNA mixture was transformed into *Saccharomyces cerevisiae* W3124 cells (MATa; ura 3-52; leu 2-3, 112; his 3-D200; pep 4-1137; prcl::HIS3; prbl::LEU2; cir+; van den Hazel *et al.*, 1992, *Eur. J. Biochem.* 207: 277-283). Cultivation, transformation
10 and media were as described by Guthrie *et al.*, 1991, *Meth. Enzymol.* Vol 194, Academic Press. The transformed cells were plated onto synthetic complete agar containing 2% glucose for 3 days at 30°C. After 3 days the colonies were replica plated to SC agar plates with 2% galactose and incubated for 4 days at 30°C. Xylanase expressing colonies were identified using a 1% agarose overlay with 0.1% AZCL-birch-xylan at pH 4.5 (Dalbøge, 2006,
15 *FEMS Microbiology Reviews* 21: 29-42). Colonies expressing xylanase activity were surrounded by a blue zone. Plasmid DNA, rescued from the positive colonies, contained a DNA insert of approximately 1.3 kb. Sequencing of the isolated gene fragment revealed a 1218 bp open reading frame encoding a polypeptide with a theoretical molecular weight of 43.0 kDa. The cDNA fragment was subcloned into the *Aspergillus* expression vector
20 pHD464 (Dalbøge and Heldt-Hansen, 1994, *Mol. Gen. Genet.* 243, 253-260) digested with *Bam* HI and *Xho* I by digesting the clone with *Bam* HI and *Xho* I and isolating the 1.2 kb cDNA insert (Christgau *et al.*, 1996, *Biochem. J.* 319: 705-712) to generate plasmid pA2X2.

The *A. aculeatus* GH10 xylanase coding sequence was PCR amplified using plasmid pA2x2 as template and primers 153505 and 153506 shown below using standard methods
25 to yield an approximately 1.2 kb fragment. The 1.2 kb fragment was digested with *Bam* HI and *Xho* I (introduced in the PCR primers) and cloned into vector pCaHj527 (WO 2004/099228). The resulting plasmid was designated pMT2155 in which the cDNA was under transcriptional control of the neutral amylase II (NA2) promoter from *A. niger* and the AMG terminator from *A. niger*.

30 Primer 153505:

5'-TCTTGGATCCACCATGGTCGGACTGCTTTCAATCACC-3' (SEQ ID NO: 89)

Primer 153506:

5'-TTAACTCGAGTCACAGACACTGCGAGTAATAGTC-3' (SEQ ID NO: 90)

Two synthetic oligonucleotide primers shown below were designed to PCR amplify
35 the *A. aculeatus* GH10 gene from plasmid pMT2155 and introduce flanking regions for insertion into expression vector pMJ09 (WO 2005/056772). Bold letters represent coding sequence and the remaining sequence is homologous to the insertion sites of pMJ09.

Forward Primer:

5'-cggactcgcaccat**gggtcggactgctttcaat**-3' (SEQ ID NO: 91)

Reverse Primer:

5'-tcgccacggagctt**atcacagacactgcgagtaat**-3' (SEQ ID NO: 92)

5 Fifty picomoles of each of the primers above were used in a PCR reaction consisting of 50 ng of pMT2155, 1 µl of a 10 mM blend of dATP, dTTP, dGTP, and dCTP, 5 µl of 10X ACCUTAQ™ DNA Polymerase Buffer, and 5 units of ACCUTAQ™ DNA polymerase, in a final volume of 50 µl. An EPPENDORF® MASTERCYCLER® 5333 was used to amplify the DNA fragment programmed for 1 cycle at 95°C for 3 minutes; and 30 cycles each at 94°C for 10 45 seconds, 55°C for 60 seconds, and 72°C for 1 minute 30 seconds. After the 30 cycles, the reaction was incubated at 72°C for 10 minutes and then cooled to 4°C until further processing.

15 The reaction products were isolated by 1% agarose gel electrophoresis using TAE buffer where a 1.2 kb product band was excised from the gel and purified using a QIAquick Gel Extraction Kit according to the manufacturer's instructions.

20 The fragment was then cloned into pMJ09 using an IN-FUSION™ Cloning Kit. The vector was digested with *Nco* I and *Pac* I and purified by agarose gel electrophoresis as described above. The 1.2 kb gene fragment and the digested vector were ligated together in a reaction resulting in the expression plasmid pSaMe-AaXYL in which transcription of the Family GH10 gene was under the control of the *T. reesei* *cbh1* promoter. The ligation reaction (50 µl) was composed of 1X IN-FUSION™ Buffer, 1X BSA, 1 µl of IN-FUSION™ enzyme (diluted 1:10), 100 ng of pAIIo2 digested with *Nco* I and *Pac* I, and 100 ng of the *A. aculeatus* GH10 xylanase purified PCR product. The reaction was incubated at room temperature for 30 minutes. One µl of the reaction was used to transform *E. coli* XL10 25 SOLOPACK® Gold competent cells according to the manufacturer. An *E. coli* transformant containing pSaMe-AaGH10 was detected by restriction enzyme digestion and plasmid DNA was prepared using a BIOROBOT® 9600. DNA sequencing of the *A. aculeatus* GH10 gene from pSaMe-AaXYL was performed using dye-terminator chemistry (Giesecke *et al.*, 1992, *supra*) and primer walking strategy.

30 Plasmids pSaMe-AaXYL encoding the *A. aculeatus* GH10 endoglucanase and pSaMe-TrBXYL encoding the *T. reesei* beta-xylosidase were co-transformed into *Trichoderma reesei* RutC30 by PEG-mediated transformation (Penttila *et al.*, 1987, *Gene* 61 155-164) to generate *T. reesei* strain SaMe-BXX13. Each plasmid contained the *A. nidulans* *amdS* gene to enable transformants to grow on acetamide as the sole nitrogen source.

35 *T. reesei* RutC30 was cultivated at 27°C and 90 rpm in 25 ml of YP medium supplemented with 2% (w/v) glucose and 10 mM uridine for 17 hours. Mycelia were collected by filtration using a Vacuum Driven Disposable Filtration System (Millipore, Bedford, MA,

USA) and washed twice with deionized water and twice with 1.2 M sorbitol. Protoplasts were generated by suspending the washed mycelia in 20 ml of 1.2 M sorbitol containing 15 mg of GLUCANEX™ (Novozymes A/S, Bagsvaerd, Denmark) per ml and 0.36 units of chitinase (Sigma Chemical Co., St. Louis, MO, USA) per ml and incubating for 15-25 minutes at 34°C with gentle shaking at 90 rpm. Protoplasts were collected by centrifuging for 7 minutes at 400 x g and washed twice with cold 1.2 M sorbitol. The protoplasts were counted using a haemocytometer and re-suspended in STC to a final concentration of 1 X 10⁸ protoplasts per ml. Excess protoplasts were stored in a Cryo 1°C Freezing Container (Nalgene, Rochester, NY, USA) at -80°C.

Approximately 4 µg of plasmids pSaMe-AaXYL and pSaMe-TRBXYL were digested with *Pme* I and added to 100 µl of protoplast solution and mixed gently, followed by 250 µl of 10 mM CaCl₂-10 mM Tris-HCl pH 7.5-60% PEG 4000, mixed, and incubated at room temperature for 30 minutes. STC (3 ml) was then added and mixed and the transformation solution was plated onto COVE plates using *Aspergillus nidulans amdS* selection. The plates were incubated at 28°C for 5-7 days. Transformants were sub-cultured onto COVE2 plates and grown at 28°C.

Over 40 transformants were subcultured onto fresh plates containing acetamide and allowed to sporulate for 7 days at 28°C.

The *Trichoderma reesei* transformants were cultivated in 125 ml baffled shake flasks containing 25 ml of CIM medium at pH 6.0 by inoculating spores of the transformants and incubating at 28°C and 200 rpm for 7 days. *Trichoderma reesei* RutC30 was run as a control. Culture broth samples were removed at day 5. One ml of each culture broth was centrifuged at 15,700 x g for 5 minutes in a microcentrifuge and the supernatants transferred to new tubes.

SDS-PAGE was performed using CRITERION® Tris-HCl (5% resolving) gels with a CRITERION® System (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Five µl of day 7 supernatants (see above) were suspended in 2X concentration of Laemmli Sample Buffer (Bio-Rad Laboratories, Inc., Hercules, CA, USA) and boiled in the presence of 5% beta-mercaptoethanol for 3 minutes. The supernatant samples were loaded onto a polyacrylamide gel and subjected to electrophoresis with 1X Tris/Glycine/SDS as running buffer (Bio-Rad Laboratories, Inc., Hercules, CA, USA). The resulting gel was stained with BIO-SAFE™ Coomassie Stain. The transformant showing the highest expression of both the *A. aculeatus* GH10 xylanase and the *T. reesei* beta-xylosidase based on the protein gel was designated *T. reesei* SaMe-BXX13.

Trichoderma reesei SaMe-BXX13 was cultivated in 500 ml baffled shake flasks containing 250 ml of CIM medium at pH 6.0 inoculated with spores of *T. reesei* SaMe-

BXX13. Shake flasks were incubated at 28°C at 200 rpm for five days. The culture broth was then filtered using an 0.22 µm EXPRESS™ Plus Membrane.

The filtered broth was concentrated and buffer exchanged using a tangential a tangential flow concentrator (Pall Filtron, Northborough, MA, USA) equipped with a 10 kDa polyethersulfone membrane (Pall Filtron, Northborough, MA, USA) to pH 4.0 with acetic acid. Sample was loaded onto a SP SEPHAROSE® column (GE Healthcare, Piscataway, NJ, USA) equilibrated in 50 mM sodium acetate pH 4.0, eluting bound proteins with a gradient of 0-1000 mM sodium chloride. Fractions were buffer exchanged into 20 mM sodium phosphate pH 7.0 using a tangential flow concentrator and applied to a PHENYL SUPEROSE™ HR 16/10 column (GE Healthcare, Piscataway, NJ, USA) equilibrated with 1.5 M (NH₄)₂SO₄-20 mM sodium phosphate pH 7.0. Bound proteins were eluted with a linear gradient over 20 column volumes from 1.5 to 0 M (NH₄)₂SO₄ in 20 mM Tris-HCl pH 7.0. The protein fractions were buffer exchanged into 20 mM TEA HCl pH 7.5 using a tangential flow concentrator. Sample was applied to a MONOQ™ HR 16/10 ion exchange column, equilibrated in 20 mM TEA HCl pH 7.5, eluting bound proteins with a gradient from 0-300 mM sodium chloride. Buffer of final protein fractions was 20 mM TEA-100 mM sodium chloride pH 7.5. Protein concentration was determined using a Microplate BCA™ Protein Assay Kit in which bovine serum albumin was used as a protein standard.

Example 11: Preparation of phosphoric acid swollen cellulose

Phosphoric acid swollen cellulose (PASC) was prepared from AVICEL® PH101 (FMC, Philadelphia, PA, USA) using the prortocol described by Zhang *et al.*, 2006, *Biomacromolecules* 7: 644–648.

Example 12: Phosphoric acid swollen cellulose (PASC) hydrolysis assay

A 1.0% slurry of PASC prepared as described in Example 11 was thoroughly resuspended by shaking, and quickly transferred to a 100 ml beaker and stirred rapidly with a magnetic stirrer. Five hundred µl aliquots of the 1.0% PASC slurry were pipetted into wells of a 2.0 ml 96-deepwell plate (Axygen, Union City, CA, USA) using a 1000 µl micropipette with a wide aperture tip (end of tip cut off about 2 mm from the base). One hundred µl of 10 mM MnSO₄-500 mM sodium acetate pH 5 were then added to each well. Two hundred µl of either deionized water or a 1.0% pyrogallol (w/w) (Sigma Chemical Co., Inc., St. Louis, Mo, USA) solution were added to each well. Enzyme mixtures were prepared and then added simultaneously to all wells in a volume of 200 µl, for a total 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 either 50°C or 65°C for approximately 3 days.

All experiments reported were performed in triplicate.

Primary analysis of the hydrolysis reactions was performed using an AGILENT® 1100 HPLC (Agilent Technologies, Inc., Santa Clara, CA, USA) with CHEMSTATION® software (Agilent Technologies, Inc., Santa Clara, CA, USA) equipped with an AMINEX™ HPX-87H column (Bio-Rad Laboratories, Inc., Hercules, CA, USA). After approximately 4 days, the deep-well plate was removed from the incubator and chilled overnight to 4°C. The plate was then mixed well by inversion and briefly centrifuged at 52 x g in a SORVALL® RT7 centrifuge for 10 seconds. Samples were then mixed by pipetting, and 200 µl from each well were transferred to a MULTISCREEN® HV (Millipore, Bedford, MA, USA) centrifuge filter plate assembly. The centrifuge filter plate assembly was centrifuged at 2000 rpm in a SORVALL® RT7 centrifuge for 20 minutes. The filtrates were transferred to a 96 well autosampler plate and diluted 1:1 with 5 mM H₂SO₄, sealed with silicon sealing mat, and inserted into an HPLC injector module (set to 4°C) for injection of 20µl onto a CATION H™ guard column connected to a 4.6 x 250 mm AMINEX® HPX-87H column followed by elution with 0.05% w/w benzoic acid in 5 mM H₂SO₄. Sugars were detected by refractive index detection with quantification by integration compared to purified sugar standards.

All HPLC data processing was performed using MICROSOFT EXCEL™ software (Microsoft, Redmond, WA, USA). Measured glucose concentrations were adjusted for the appropriate dilution factor. In this assay only glucose was measured since beta-glucosidase was at high levels in all samples but the controls. Percent relative conversion was calculated using the following Equation:

$$\% \text{ conversion} = [\text{sample glucose concentration}] / [\text{glucose concentration of limit digest}] \times 100$$

In order to calculate % conversion, a 100% conversion point was set based on a cellulase control of 100 mg of *Trichoderma reesei* cellulase per gram cellulose (CELLUCLAST PLUS™, Novozymes A/S, Bagsvaerd, Denmark), and all values were divided by this number and then multiplied by 100. Triplicate data points were averaged and standard deviation was calculated.

Example 13: Effect of the addition of *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V on conversion of phosphoric acid swollen cellulose by *Aspergillus fumigatus* beta-glucosidase in the presence of pyrogallol at 50°C and 65°C

A. fumigatus GH61B wild-type polypeptide and *A. fumigatus* GH61B variant L90V I75V + F77L + F179I + I181L + I183V were evaluated for their ability to enhance the hydrolysis of phosphoric acid swollen cellulose by *A. fumigatus* CEL3A beta-glucosidase in the presence of pyrogallol. The phosphoric acid swollen cellulose hydrolysis assay was performed as described in Example 12.

The conversion of phosphoric acid swollen cellulose (0.5% w/w) by the combination of pyrogallol (0.2% w/w) and *A. fumigatus* CEL3A beta-glucosidase (5 mg protein per g cellulose); the combination of (0.2% w/w) pyrogallol, *A. fumigatus* GH61B wild-type polypeptide (10 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (5 mg protein per g cellulose); and the combination of pyrogallol (0.2% w/w), *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (10 mg protein per g cellulose) and *A. fumigatus* CEL3A beta-glucosidase (5 mg protein per g cellulose) was determined according to the assay described in Example 12. Data were collected and analyzed, as described in Example 12, after 72 hours of incubation at either 50°C or 65°C. The results are shown in Figure 2.

The combination of pyrogallol (0.2% w/w) and *A. fumigatus* CEL3A beta-glucosidase (5 mg protein per g cellulose) resulted in conversion of phosphoric acid swollen cellulose of $2.1 \pm 0.3\%$ and $2.2 \pm 1.0\%$ at 50°C and 65°C, respectively. Addition of *A. fumigatus* GH61B wild-type polypeptide (10 mg protein per g cellulose) to the combination of pyrogallol (0.2% w/w) and *A. fumigatus* CEL3A beta-glucosidase (5 mg protein per g cellulose) resulted in the conversion of phosphoric acid swollen cellulose of $23.8 \pm 0.6\%$ and $16.8 \pm 0.4\%$ at 50°C and 65°C, respectively. Addition of *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (10 mg protein per g cellulose) to the combination of pyrogallol (0.2% w/w) and *A. fumigatus* CEL3A beta-glucosidase (5 mg protein per g cellulose) resulted in the conversion of phosphoric acid swollen cellulose of $29.0 \pm 1.0\%$ and $21.1 \pm 0.3\%$ at 50°C and 65°C, respectively.

Example 14: Pretreated corn stover hydrolysis assay

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 approximately 59% cellulose, 5% hemicellulose and 28% 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.

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 PCS per ml of 50 mM sodium acetate pH 5.0 buffer containing 1 mM manganese sulfate and various protein loadings of the cellulase composition (expressed as mg protein per gram of cellulose). Enzyme mixtures were prepared and then added simultaneously to

all wells in a volume of 100 µl, for a final volume of 1 ml in each reaction. The plate was then sealed using an ALPS-300™ plate heat sealer, mixed thoroughly, and incubated at 50°C, 55°C, 60°C, and/or 65°C for 72 hours. All experiments were performed in triplicate.

5 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 sugary 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 by elution with 0.05% w/w benzoic acid-0.005 M H₂SO₄ at a flow rate of 0.6 ml per minute at 65°C, and quantitation by integration of glucose and cellobiose signal from refractive index detection (CHEMSTATION®, AGILENT® 1100
10 HPLC, Agilent Technologies, Santa Clara, CA, USA) calibrated by pure sugar samples. The resultant equivalents were used to calculate the percentage of cellulose conversion for each reaction.

All HPLC data processing was performed using MICROSOFT EXCEL™ software.
15 Measured sugar concentrations were adjusted for the appropriate dilution factor. Glucose and cellobiose were measured individually. However, 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 cellulose conversion was calculated using the following equation:

$$20 \quad \% \text{ conversion} = \frac{[\text{sample glucose concentration}]}{[\text{glucose concentration in a limit digest}]} \times 100$$

In order to calculate % conversion, a 100% conversion point was set based on a cellulase control of 100 mg of the cellulase composition per gram cellulose (CELLUCLAST PLUS™, Novozymes A/S, Bagsvaerd, Denmark), and all values were divided by this number
25 and then multiplied by 100. Triplicate data points were averaged and standard deviation was calculated.

Example 15: Effect of the addition of *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V on conversion of PCS by a *Trichoderma reesei* cellulase composition and *Aspergillus fumigatus* CEL3A beta-glucosidase at 50°C and 55°C
30

A cellulase composition from *Trichoderma reesei* strain 981-O-8 (D4) was prepared according to the fermentation protocol described below. *T. reesei* strain 981-O-8 (D4) is a mutagenized strain of *Trichoderma reesei* RutC30 (ATCC 56765; Montenecourt and Eveleigh, 1979, *Adv. Chem. Ser.* 181: 289-301). The cellulase composition is similar to
35 CELLUCLAST (Novozymes A/S, Bagsvaerd, Denmark). The cellulase composition was supplemented with *Aspergillus fumigatus* beta-glucosidase. The cellulase composition supplemented with *A. fumigatus* beta-glucosidase is designated in the Examples as

"*Trichoderma reesei* cellulase composition".

Shake flask medium was composed of 20 g of dextrose, 10 g of corn steep solids, 1.45 g of $(\text{NH}_4)_2\text{SO}_4$, 2.08 g of KH_2PO_4 , 0.36 g of CaCl_2 , 0.42 g of $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.42 ml of trace metals solution, and deionized water to 1 liter. Trace metals solution was composed of 216 g of $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, 58 g of $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 27 g of $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 10 g of $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 2.4 g of H_3BO_3 , 336 g of citric acid, and deionized water to 1 liter.

One hundred ml of shake flask medium was added to a 500 ml shake flask. The shake flask was inoculated with two plugs from a solid plate culture and incubated at 28°C on an orbital shaker at 200 rpm for 48 hours. Fifty ml of the shake flask broth was used to inoculate a 2 liter fermentation vessel.

Fermentation batch medium was composed of 30 g of cellulose, 4 g of dextrose, 10 g of corn steep solids, 3.8 g of $(\text{NH}_4)_2\text{SO}_4$, 2.8 g of KH_2PO_4 , 2.64 g of CaCl_2 , 1.63 g of $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1.8 ml of anti-foam, 0.66 ml of trace metals solution, and deionized water to 1 liter. Trace metals solution was composed of 216 g of $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, 58 g of $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 27 g of $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 10 g of $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 2.4 g of H_3BO_3 , 336 g of citric acid, and deionized water to 1 liter. Fermentation feed medium was composed of dextrose.

A total of 1.8 liters of the fermentation batch medium was added to an Applikon Biotechnology three liter glass jacketed fermentor. Fermentation feed medium was dosed at a rate of 0 to 4 g/l/hr for a period of 185 hours. The fermentation vessel was maintained at a temperature of 28°C and pH was controlled using an Applikon 1030 control system to a set-point of 4.5 +/- 0.1. Air was added to the vessel at a rate of 1 vvm and the broth was agitated by 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.

A. fumigatus GH61B wild-type polypeptide and *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V were evaluated for their ability to enhance the hydrolysis of PCS by the combination of the *Trichoderma reesei* cellulase composition and *A. fumigatus* CEL3A beta-glucosidase at either 50°C or 55°C. The PCS hydrolysis assay was performed as described in Example 14.

The conversion of PCS by the combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose) and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose); the combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B wild-type polypeptide (0.3 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose); the combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B wild-type polypeptide (0.45 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose); the combination of the *T. reesei* cellulase

composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B wild-type polypeptide (0.6 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose); the combination of a *Trichoderma* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B wild-type polypeptide (0.75 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose); the combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.3 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose); the combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.45 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose); the combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.6 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose); and the combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.75 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose) was determined according to Example 15. Data were collected and analyzed, as described in Example 15, after 72 hours of incubation at either 50°C or 55°C. Results for 50°C and 55°C are shown in Figures 3A and 3B, respectively.

The combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose) and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose) resulted in conversion of PCS of $53.5 \pm 0.2\%$ and $52.2 \pm 0.3\%$ at 50°C and 55°C, respectively.

The combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B wild-type polypeptide (0.3 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose) resulted in conversion of PCS of $62.6 \pm 0.1\%$ and $61.5 \pm 0.4\%$ at 50°C and 55°C, respectively. The combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B wild-type polypeptide (0.45 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose) resulted in conversion of PCS of $64.3 \pm 0.1\%$ and $63.4 \pm 0.1\%$ at 50°C and 55°C, respectively.

The combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B wild-type polypeptide (0.6 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose) resulted in conversion of PCS of $65.0 \pm 0.5\%$ and $64.5 \pm 0.6\%$ at 50°C and 55°C, respectively. The combination of the

T. reesei cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B wild-type polypeptide (0.75 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose) resulted in conversion of PCS of $66.4 \pm 0.2\%$ and $66.1 \pm 0.4\%$ at 50°C and 55°C, respectively.

5 The combination of the *T. reesei* cellulase composition (2.7 mg protein per gram cellulose), *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.3 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose) resulted in conversion of PCS of $63.3 \pm 0.3\%$ and $62.3 \pm 0.4\%$ at 50°C and 55°C, respectively. The combination of the *T. reesei* cellulase composition (2.7 mg protein per
10 gram cellulose), *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.45 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose) resulted in conversion of PCS of $65.0 \pm 0.2\%$ and $64.5 \pm 0.2\%$ at 50°C and 55°C, respectively. The combination of the *T. reesei* cellulase composition (2.7 mg protein per
15 gram cellulose), *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.6 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose) resulted in conversion of PCS of $65.7 \pm 0.2\%$ and $65.7 \pm 0.5\%$ at 50°C and 55°C, respectively. The combination of the *T. reesei* cellulase composition (2.7 mg protein per
20 gram cellulose), *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.75 mg protein per g cellulose), and *A. fumigatus* CEL3A beta-glucosidase (0.3 mg protein per g cellulose) resulted in conversion of PCS of $67.0 \pm 0.2\%$ and $66.9 \pm 0.3\%$ at 50°C and 55°C, respectively.

Example 16: Effect of the addition of *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V on the conversion of PCS by a high-temperature cellulase composition at 50°C, 55°C, 60°C, and 65°C

25 A mixture of cellulase enzymes, designated "high temperature cellulase composition", was prepared by mixing prepared enzyme components in the following ratios (of total protein): 43.5% *Aspergillus fumigatus* cellobiohydrolase I, 29.4% *Aspergillus fumigatus* cellobiohydrolase II, 5.9% *Aspergillus fumigatus* beta-glucosidase, 5.9%
30 *Aspergillus fumigatus* GH10 xylanase 3, 3.5% *Trichoderma reesei* beta-xylosidase, and 11.8% *Trichoderma reesei* CEL5A endoglucanase II. The composition when loaded in assay at 3 mg total protein per gram cellulose had the following enzyme loadings per gram cellulose: 1.31 mg *A. fumigatus* cellobiohydrolase I per gram cellulose, 0.88 mg *A. fumigatus* cellobiohydrolase II per gram cellulose, 0.18 mg *A. fumigatus* beta-glucosidase per gram
35 cellulose, 0.18 mg *Aspergillus fumigatus* GH10 xylanase 3 per gram cellulose, 0.11 mg *T. reesei* beta-xylosidase per gram cellulose, and 0.35 mg *T. reesei* CEL5A endoglucanase II

per gram cellulose.

A. fumigatus GH61B wild-type polypeptide and *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V were evaluated for their ability to enhance the hydrolysis of PCS by the high temperature cellulase composition at 50°C, 55°C, 60°C, and 65°C. The pretreated corn stover hydrolysis assay was performed as described in Example 14.

The conversion of pretreated corn stover by the high temperature cellulase composition (3 mg protein per g cellulose); the combination of the high temperature cellulase composition (3 mg protein per g cellulose) and *A. fumigatus* GH61B wild-type polypeptide (0.45 mg protein per g cellulose); the combination of the high temperature cellulase composition (3 mg protein per g cellulose) and *A. fumigatus* GH61B wild-type polypeptide (0.75 mg protein per g cellulose); the combination of the high temperature cellulase composition (3 mg protein per g cellulose) and *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.45 mg protein per g cellulose); and the combination of the high temperature cellulase composition (3 mg protein per g cellulose) and *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.75 mg protein per g cellulose) was assayed as described in Example 14. Data were collected and analyzed, as described in Example 14, after 72 hours of incubation at 50°C, 55°C, 60°C, and 65°C. Results for 0.45 mg addition of GH61 polypeptide per gram cellulose and 0.75 mg addition of GH61 polypeptide per gram cellulose are shown in Figures 4A and 4B, respectively.

The high temperature cellulase composition (3 mg protein per g cellulose) resulted in conversion of pretreated corn stover of $48.5 \pm 0.3\%$, $52.0 \pm 0.1\%$, $43.7 \pm 0.2\%$, and $36.4 \pm 0.1\%$ at 50°C, 55°C, 60°C, and 65°C, respectively.

The combination of the high temperature cellulase composition (3 mg protein per g cellulose) and *A. fumigatus* GH61B wild-type polypeptide (0.45 mg protein per g cellulose) resulted in conversion of PCS of $60.4 \pm 0.2\%$, $65.8 \pm 0.1\%$, $55.8 \pm 0.1\%$, and $45.2 \pm 0.2\%$ at 50°C, 55°C, 60°C, or 65°C, respectively. The combination of the high temperature cellulase composition (3 mg protein per g cellulose) and *A. fumigatus* GH61B wild-type polypeptide (0.75 mg protein per g cellulose) resulted in conversion of PCS of $61.9 \pm 0.1\%$, $67.6 \pm 0.2\%$, $57.8 \pm 0.3\%$, and $45.5 \pm 0.7\%$ at 50°C, 55°C, 60°C, and 65°C, respectively.

The combination of the high temperature cellulase composition (3 mg protein per g cellulose) and *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.45 mg protein per g cellulose) resulted in conversion of PCS of $51.4 \pm 0.7\%$, $66.9 \pm 0.5\%$, $56.8 \pm 0.4\%$, and $46.2 \pm 0.1\%$ at 50°C, 55°C, 60°C, and 65°C, respectively.

The combination of the high temperature cellulase composition (3 mg protein per g cellulose) and *A. fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V (0.75 mg protein per g cellulose) resulted in conversion of PCS of $62.3 \pm 0.2\%$, $68.3 \pm 0.2\%$, $58.8 \pm$

0.4%, and $46.2 \pm 1.1\%$ at 50°C, 55°C, 60°C, and 65°C, respectively.

Example 17: Determination of Td (denaturation temperature) of the *Aspergillus fumigatus* wild-type GH61B polypeptide and the *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V by differential scanning calorimetry

The thermostabilities of the *A. fumigatus* wild-type GH61B polypeptide and the *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V GH61 polypeptide were determined by Differential Scanning Calorimetry (DSC) using a VP-Capillary Differential Scanning Calorimeter with autosampler (MicroCal Inc., GE Health Care, Piscataway, NJ, USA). The thermal denaturation temperature, Td (°C), was taken as the top of denaturation peak (major endothermic peak) in thermograms (Cp vs. T) obtained after heating the enzyme solutions in 50 mM sodium acetate pH 5.0 with 100 ppm TRITON® X100 added at a constant programmed heating rate. Approximately 0.4 ml of sample and reference-solutions were stored at 10°C prior to loading of samples into the calorimeter. Sample and reference (reference: buffer without enzyme) solutions were automatically loaded into the DSC and thermally pre-equilibrated for 20 minutes at 20°C before the DSC scan was performed from 20°C to 90°C at a scan rate of 200 K/hr. Denaturation temperatures were determined at an accuracy of approximately +/- 1°C. The results are shown in Figures 5A and 5B.

By differential scanning calorimetry, the *A. fumigatus* wild-type GH61B polypeptide has a Td of approximately 68°C at pH 5 (Figure 5A), while the *Aspergillus fumigatus* GH61B variant I75V + F77L + F179I + I181L + I183V GH61 polypeptide has a Td of approximately 74°C at pH 5 (Figure 5B).

The present invention is further described by the following numbered paragraphs:

[1] A variant, comprising a substitution at one or more positions corresponding to positions 75, 77, 179, 181, and 183 of the mature polypeptide of SEQ ID NO: 2, wherein the variant has cellulolytic enhancing activity.

[2] The variant of paragraph 1, which is a variant of a parent polypeptide having cellulolytic enhancing activity selected from the group consisting of: (a) a polypeptide having at least 60% sequence identity to the mature polypeptide of SEQ ID NO: 2; (b) a polypeptide encoded by a polynucleotide that hybridizes under at least low 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); (c) a polypeptide encoded by a polynucleotide having at least 60% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and (d) a fragment of the mature polypeptide of SEQ ID NO: 2, which has cellulolytic enhancing activity.

[3] The variant of paragraph 2, wherein the parent polypeptide having cellulolytic enhancing activity has at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, 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% sequence identity to the mature polypeptide of SEQ ID NO: 2.

[4] The variant of paragraph 2, wherein the parent polypeptide having cellulolytic enhancing activity is encoded by a polynucleotide that hybridizes under 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).

[5] The variant of paragraph 2, wherein the parent polypeptide having cellulolytic enhancing activity is encoded by a polynucleotide having at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, 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% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof.

[6] The variant of paragraph 2, wherein the parent polypeptide having cellulolytic enhancing activity comprises or consists of the mature polypeptide of SEQ ID NO: 2.

[7] The variant of paragraph 2, wherein the parent polypeptide having cellulolytic enhancing activity is a fragment of the mature polypeptide of SEQ ID NO: 2, wherein the fragment has cellulolytic enhancing activity.

[8] The variant of paragraph 2, which has at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, 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%, or at least 99%, but less than 100%, sequence identity to the amino acid sequence of the parent polypeptide having cellulolytic enhancing activity.

[9] The variant of any of paragraphs 1-8, which has at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, 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%, or at least 99%, but less than 100% sequence identity to the mature polypeptide of SEQ ID NO: 2.

[10] The variant of any of paragraphs 1-9, wherein the variant consists of 200 to 220 amino acids, e.g., 205 to 220, 210 to 220, and 215 to 220 amino acids.

[11] The variant of any of paragraphs 1-10, wherein the number of substitutions is 1-5, e.g., such as 1, 2, 3, 4, or 5 substitutions.

5 [12] The variant of any of paragraphs 1-11, which comprises a substitution at a position corresponding to position 75.

[13] The variant of paragraph 12, wherein the substitution is Val.

[14] The variant of any of paragraphs 1-13, which comprises a substitution at a position corresponding to position 77.

10 [15] The variant of paragraph 14, wherein the substitution is Ser.

[16] The variant of any of paragraphs 1-15, which comprises a substitution at a position corresponding to position 179.

[17] The variant of paragraph 16, wherein the substitution is Leu.

15 [18] The variant of any of paragraphs 1-17, which comprises a substitution at a position corresponding to position 181.

[19] The variant of paragraph 18, wherein the substitution is Trp.

[20] The variant of any of paragraphs 1-17, which comprises a substitution at a position corresponding to position 183.

[21] The variant of paragraph 20, wherein the substitution is Val.

20 [22] The variant of any of paragraphs 1-21, which comprises a substitution at two positions corresponding to any of positions 75, 77, 179, 181, and 183.

[23] The variant of any of paragraphs 1-21, which comprises a substitution at three positions corresponding to any of positions 75, 77, 179, 181, and 183.

25 [24] The variant of any of paragraphs 1-21, which comprises a substitution at four positions corresponding to any of positions 75, 77, 179, 181, and 183.

[25] The variant of any of paragraphs 1-21, which comprises a substitution at each position corresponding to positions 75, 77, 179, 181, and 183.

30 [26] The variant of any of paragraphs 1-25, which comprises or consists of one or more substitutions selected from the group consisting of I75V, F77L, F179I, I181L, and I183V.

[27] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F77L of the mature polypeptide of SEQ ID NO: 2.

[28] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F179I of the mature polypeptide of SEQ ID NO: 2.

35 [29] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + I181L of the mature polypeptide of SEQ ID NO: 2.

[30] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + I183V of the mature polypeptide of SEQ ID NO: 2.

[31] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions F77L + F179I of the mature polypeptide of SEQ ID NO: 2.

5 [32] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions F77L + I181L of the mature polypeptide of SEQ ID NO: 2.

[33] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions F77L + I183V of the mature polypeptide of SEQ ID NO: 2.

10 [34] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions F179I + I181L of the mature polypeptide of SEQ ID NO: 2.

[35] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions F179I + I183V of the mature polypeptide of SEQ ID NO: 2.

[36] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

15 [37] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F77L + F179I of the mature polypeptide of SEQ ID NO: 2.

[38] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F77L + I181L of the mature polypeptide of SEQ ID NO: 2.

20 [39] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F77L + I183V of the mature polypeptide of SEQ ID NO: 2.

[40] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F179I + I181L of the mature polypeptide of SEQ ID NO: 2.

[41] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F179I + I183V of the mature polypeptide of SEQ ID NO: 2.

25 [42] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

[43] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions F77L + F179I + I181L of the mature polypeptide of SEQ ID NO: 2.

30 [44] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions F77L + F179I + I183V of the mature polypeptide of SEQ ID NO: 2.

[45] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions F179I + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

[46] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions F77L + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

35 [47] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F77L + F179I + I181L of the mature polypeptide of SEQ ID NO: 2.

[48] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F179I + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

[49] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F77L + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

5 [50] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F77L + F179I + I183V of the mature polypeptide of SEQ ID NO: 2.

[51] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions F77L + F179I + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

10 [52] The variant of any of paragraphs 1-25, which comprises or consists of the substitutions I75V + F77L + F179I + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

[53] An isolated polynucleotide encoding the variant of any of paragraphs 1-52.

[54] A nucleic acid construct comprising the polynucleotide of paragraph 53.

[55] An expression vector comprising the polynucleotide of paragraph 53.

15 [56] A host cell comprising the polynucleotide of paragraph 53.

[57] A method of producing a variant, comprising: (a) cultivating the host cell of paragraph 56 under conditions suitable for the expression of the variant; and (b) recovering the variant.

20 [58] A transgenic plant, plant part or plant cell transformed with the polynucleotide of paragraph 53.

[59] A method of producing the variant of any of paragraphs 1-52, comprising: (a) cultivating a transgenic plant or a plant cell comprising a polynucleotide encoding the variant under conditions conducive for production of the variant; and (b) recovering the variant.

25 [60] A method for obtaining a variant, comprising introducing into a parent polypeptide a substitution at one or more positions corresponding to positions 75, 77, 179, 181, and 183 of the mature polypeptide of SEQ ID NO: 2, wherein the variant has cellulolytic enhancing activity; and recovering the variant.

30 [61] A method for degrading or converting a cellulosic material, comprising: treating the cellulosic material with an enzyme composition in the presence of the variant of any of paragraphs 1-52.

[62] The method of paragraph 61, wherein the cellulosic material is pretreated.

[63] The method of of paragraph 61 or 62, further comprising recovering the degraded cellulosic material.

35 [64] The method of any of paragraphs 61-63, wherein the enzyme composition comprises one or more enzymes selected from the group consisting of a cellulase, a polypeptide having cellulolytic enhancing activity, a hemicellulase, an expansin, an esterase, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

[65] The method of paragraph 64, wherein the cellulase is one or more enzymes selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase.

5 [66] The method of paragraph 64, wherein the hemicellulase is one or more enzymes selected from the group consisting of a xylanase, an acetylxylan esterase, a feruloyl esterase, an arabinofuranosidase, a xylosidase, and a glucuronidase.

[67] The method of any of paragraphs 61-66, wherein the degraded cellulosic material is a sugar.

10 [68] The method of paragraph 64, wherein the sugar is selected from the group consisting of glucose, xylose, mannose, galactose, and arabinose.

[69] A method for producing a fermentation product, comprising: (a) saccharifying a cellulosic material with an enzyme composition in the presence of the variant of any of paragraphs 1-52; (b) fermenting the saccharified cellulosic material with one or more fermenting microorganisms to produce the fermentation product; and (c) recovering the
15 fermentation product from the fermentation.

[70] The method of paragraph 69, wherein the cellulosic material is pretreated.

[71] The method of paragraph 69 or 70, wherein the enzyme composition comprises one or more enzymes selected from the group consisting of a cellulase, a polypeptide having cellulolytic enhancing activity, a hemicellulase, an expansin, an esterase, a laccase, a
20 ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

[72] The method of paragraph 71, wherein the cellulase is one or more enzymes selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase.

25 [73] The method of paragraph 71, wherein the hemicellulase is one or more enzymes selected from the group consisting of a xylanase, an acetylxylan esterase, a feruloyl esterase, an arabinofuranosidase, a xylosidase, and a glucuronidase.

[74] The method of any of paragraphs 69-73, wherein steps (a) and (b) are performed simultaneously in a simultaneous saccharification and fermentation.

30 [75] The method of any of paragraphs 69-74, wherein the fermentation product is an alcohol, an organic acid, a ketone, an amino acid, or a gas.

[76] A method of fermenting a cellulosic material, comprising: fermenting the cellulosic material with one or more fermenting microorganisms, wherein the cellulosic material is saccharified with an enzyme composition in the presence of the variant of any of paragraphs 1-52.

35 [77] The method of paragraph 76, wherein the cellulosic material is pretreated before saccharification.

[78] The method of paragraph 76 or 77, wherein the enzyme composition comprises

one or more enzymes selected from the group consisting of a cellulase, a polypeptide having cellulolytic enhancing activity, a hemicellulase, an expansin, an esterase, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

5 [79] The method of paragraph 78, wherein the cellulase is one or more enzymes selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase.

[80] The method of paragraph 78, wherein the hemicellulase is one or more enzymes selected from the group consisting of a xylanase, an acetylxylan esterase, a feruloyl esterase, an arabinofuranosidase, a xylosidase, and a glucuronidase.

10 [81] The method of any of paragraphs 76-80, wherein the fermenting of the cellulosic material produces a fermentation product.

[82] The method of paragraph 81, further comprising recovering the fermentation product from the fermentation.

15 [83] The method of any of paragraphs 81 or 82, wherein the fermentation product is an alcohol, an organic acid, a ketone, an amino acid, or a gas.

[84] A detergent composition comprising the variant of any of paragraphs 1-52 and a surfactant.

[85] A whole broth formulation or cell culture composition comprising the polypeptide of any of paragraphs 1-52.

20

The invention described and claimed herein is not to be limited in scope by the specific aspects herein disclosed, since these aspects are intended as illustrations of several aspects of the invention. Any equivalent aspects are intended to be within the scope of this invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims. In the case of conflict, the present disclosure including definitions will control.

25

Claims

What is claimed is:

- 5 1. A variant, comprising a substitution at one or more positions corresponding to positions 75, 77, 179, 181, and 183 of the mature polypeptide of SEQ ID NO: 2, wherein the variant has cellulolytic enhancing activity.
2. The variant of claim 1, which is a variant of a parent polypeptide having cellulolytic
10 enhancing activity selected from the group consisting of:
- (a) a polypeptide having at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, 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
15 99%, or 100% sequence identity to the mature polypeptide of SEQ ID NO: 2;
- (b) a polypeptide encoded by a polynucleotide that hybridizes under at least low stringency conditions, *e.g.*, 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
20 full-length complement of (i) or (ii);
- (c) a polypeptide encoded by a polynucleotide having at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, 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
25 97%, at least 98%, at least 99%, or 100% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and
- (d) a fragment of the mature polypeptide of SEQ ID NO: 2, which has cellulolytic enhancing activity.
- 30 3. The variant of claim 2, which has at least 60%, *e.g.*, at least 65%, at least 70%, at least 75%, 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%, or at least 99%, but less than 100%, sequence identity to the amino acid sequence of the parent
35 polypeptide having cellulolytic enhancing activity.

4. The variant of any of claims 1-3, which has at least 60%, e.g., at least 65%, at least 70%, at least 75%, 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%, or at least 99%, but less than 100% sequence identity to the mature polypeptide of SEQ ID NO: 2.

5. The variant of any of claims 1-4, which comprises or consists of one or more substitutions selected from the group consisting of I75V, F77L, F179I, I181L, and I183V.

6. The variant of any of claims 1-4, which comprises or consists of the substitutions I75V + F77L + F179I + I181L + I183V of the mature polypeptide of SEQ ID NO: 2.

7. An isolated polynucleotide encoding the variant of any of claims 1-6.

8. A host cell comprising the polynucleotide of claim 7.

9. A method of producing a variant, comprising:

- (a) cultivating the host cell of claim 8 under conditions suitable for the expression of the variant; and
- (b) recovering the variant.

10. A transgenic plant, plant part or plant cell transformed with the polynucleotide of claim 7.

11. A method of producing the variant of any of claims 1-6, comprising:

- (a) cultivating a transgenic plant or a plant cell comprising a polynucleotide encoding the variant under conditions conducive for production of the variant; and
- (b) recovering the variant.

12. A method for obtaining a variant, comprising introducing into a parent polypeptide a substitution at one or more positions corresponding to positions 75, 77, 179, 181, and 183 of the mature polypeptide of SEQ ID NO: 2, wherein the variant has cellulolytic enhancing activity; and recovering the variant.

13. A method for degrading or converting a cellulosic material, comprising: treating the cellulosic material with an enzyme composition in the presence of the variant of any of

claims 1-6.

14. The method of of claim 13, further comprising recovering the degraded cellulosic material.

5

15. The method of claim 13 or 14, wherein the enzyme composition comprises one or more enzymes selected from the group consisting of a cellulase, a polypeptide having cellulolytic enhancing activity, a hemicellulase, an expansin, an esterase, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

10

16. A method for producing a fermentation product, comprising:

(a) saccharifying a cellulosic material with an enzyme composition in the presence of the variant of any of claims 1-6;

15

(b) fermenting the saccharified cellulosic material with one or more fermenting microorganisms to produce the fermentation product; and

(c) recovering the fermentation product from the fermentation.

17. The method of claim 16, wherein the enzyme composition comprises one or more enzymes selected from the group consisting of a cellulase, a polypeptide having cellulolytic enhancing activity, a hemicellulase, an expansin, an esterase, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

20

18. A method of fermenting a cellulosic material, comprising: fermenting the cellulosic material with one or more fermenting microorganisms, wherein the cellulosic material is saccharified with an enzyme composition in the presence of the variant of any of claims 1-6.

25

19. The method of claim 18, wherein the enzyme composition comprises one or more enzymes selected from the group consisting of a cellulase, a polypeptide having cellulolytic enhancing activity, a hemicellulase, an expansin, an esterase, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

30

20. The method of claim 18 or 19, wherein the fermenting of the cellulosic material produces a fermentation product.

35

21. The method of claim 20, further comprising recovering the fermentation product from the fermentation.

22. A detergent composition comprising the variant of any of claims 1-6 and a surfactant.

23. A whole broth formulation or cell culture composition comprising the variant of any of claims 1-6.

1/6

atgagattcggttggctcgaggtggccgctctgacggccgcttctgtagccaatgccaggtttgtgat 69
M R F G W L E V A A L T A A S V A N A Q
gctttcccgcctattgtttcggatatagttgacaatagtcgatgaaataatcaggaattggctttctctc 138
E L A F S
caccattctacccttcgccttgggctgatggccagggagagtgggcagatgcccatcgacgcgcgctcg 207
P F Y P S P W A D G Q G E W A D A H R R A V
agatcgtttctcagatgacactggcggagaaggttaaccttacaacgggtactgggtgggttgcgactt 276
E I V S Q M T L A E K V N L T T G T G
ttttgttgacagtgagctttcttactgaccatctacacagatgggaaatggaccgatgcgtcggtcaa 345
W E M D R C V G Q
accggcagcgttcccaggtgaagcttgcaattctgcaacaacgtgcaagtgtagtgtctaaaacgcggtg 414
T G S V P R
gtgcagacttgggtatcaactggggcttttggggccaggttccccttgggtatccgtttctgtgagct 483
L G I N W G L C G Q D S P L G I R F S
ataccgcgggagtgctttcagtccttgtattatgtgctgatgattgtctctgtatagctgacctcaactc 552
D L N S
cgcttccctgctggtactaatgtcgccgcgacatgggacaagacactcgctaccttctgtggcaaggc 621
A F P A G T N V A A T W D K T L A Y L R G K A
catgggtgaggaattcaacgacaagggcgtggacattttgctggggcctgctgctggtcctctcggcaa 690
M G E E F N D K G V D I L L G P A A G P L G K
ataccgcggacggcggcagaatctgggaaggcttctcctgatccggttctcactggtgtacttttcgc 759
Y P D G G R I W E G F S P D P V L T G V L F A
cgaaactatcaagggtatccaagacgcgggtgtgattgctactgccaaagcattacattctgaatgaaca 828
E T I K G I Q D A G V I A T A K H Y I L N E Q
ggagcatttccgacaggttggcggaggccagggatattggttacaacatcacggagacgatcagctcaa 897
E H F R Q V G E A Q G Y G Y N I T E T I S S N
cgtggatgacaagaccatgcacgagttgtaccttgggtgagtagttgacactgcaaagaggaccttga 966
V D D K T M H E L Y L W
ttgatttgactgacctggaatgcaggcccttggcagatgctgtgctgctgtaagattttccgtagacttg 1035
P F A D A V R A
acctcgcgacgaagaaatcgctgacgaaccatcgtagctggcggttggcgctgtcatgtgttctacaat 1104
G V G A V M C S Y N
caaatcaacaacagctacggttgtcaaaacagtcaaaactctcaacaagctcctcaaggctgagctgggc 1173
Q I N N S Y G C Q N S Q T L N K L L K A E L G
ttccaaggcttctgctcatgagtgactggagcgtcaccacagcgggtgtcggcgctgccctcgtgggttg 1242
F Q G F V M S D W S A H H S G V G A A L A G L
gatatgtcgatgcctggagacatttcccttcgacgacggactctccttctggggcacgaacctaacgtc 1311
D M S M P G D I S F D D G L S F W G T N L T V
agtgttcttaacggcaccggttccagcctggcgctgctgatgacatggctgttcgtatcatgaccgcgtac 1380
S V L N G T V P A W R V D D M A V R I M T A Y
tacaaggttggctgacgcttctcgattccccctaacttcagctcctggaccgggatgagtagcggc 1449
Y K V G R D R L R I P P N F S S W T R D E Y G
tgggagcattctgctgtctccgagggagcctggaccaaggtgaacgacttctgtaaatgtgcagcgcagt 1518
W E H S A V S E G A W T K V N D F V N V Q R S
cactctcagatcatccgtgagattgggtgcccgtagtagctcttgaagaacacgggtgctcttctcct 1587
H S Q I I R E I G A A S T V L L K N T G A L P
ttgaccggcaaggaggttaaagtgggtgttctcggtgaagacgctggttccaaccctgggggtgctaac 1656
L T G K E V K V G V L G E D A G S N P W G A N
ggctgccccgaccgcggctgtgataacggcactcttgcctatggcctggggtagtggtactgccaacttc 1725
G C P D R G C D N G T L A M A W G S G T A N F
ccttaccttgtcaccctcgagcaggctatccagcggagaggtcatcagcaacggcggcaatgtctttgct 1794
P Y L V T P E Q A I Q R E V I S N G G N V F A
gtgactgataacggggctctcagccagatggcagatggttgcattcctcaatccaggtgagtgccggctctt 1863
V T D N G A L S Q M A D V A S Q S S

Fig. 1A

2/6

```

agaaaaagaacggttctctgaatgaagttttttaaccattgCGAACAGCGTGTCTTTGGTGTGTTGTC AAC 1932
                                     V S L V F V N
gccgactctggagaggggtttcatcagtgTGCAGCGGCAACGAGGGTGACCGCAAAAATCTCACTCTGTGG 2001
A D S G E G F I S V D G N E G D R K N L T L W
aagaacggcgaggccgTcattgacactgTtTgtcagccactgcaacaacacgattgtggTtattcacagt 2070
K N G E A V I D T V V S H C N N T I V V I H S
gttgggcccgtcttgatcgaccgTggtatgataacCCCAACGTCACTGCCATCATCTGGGCCGGCTTG 2139
V G P V L I D R W Y D N P N V T A I I W A G L
cccggTcaggagagTggcaactccctggtcgacgTgctctatggccgCgtcaacCCcagcGccaagacc 2208
P G Q E S G N S L V D V L Y G R V N P S A K T
ccgTtCACctggggcaagactcgggagTcttacggggctcccttgctcaccgagcctaacaatgcaat 2277
P F T W G K T R E S Y G A P L L T E P N N G N
ggTgctcccaggatgatttcaacgagggcgTcttcattgactaccgTcactttgacaagcgcaatgag 2346
G A P Q D D F N E G V F I D Y R H F D K R N E
acccccatttatgagTttggccatggcttgagctacaccacctttggttactctcaccttcgggTtcag 2415
T P I Y E F G H G L S Y T T F G Y S H L R V Q
gccctcaatagTtTcgagTtTcggcatatgtcccgactagcggagagaccaagcctgcgccaacctatggt 2484
A L N S S S S A Y V P T S G E T K P A P T Y G
gagatcggtagTgcccgcgactacctgtatcccgagggTctcaaaagaattaccaagTttatttaccct 2553
E I G S A A D Y L Y P E G L K R I T K F I Y P
tggtcaactcgaccgacctcgaggattcttctgacgacCCgaactacggctgggaggactcggagtac 2622
W L N S T D L E D S S D D P N Y G W E D S E Y
attcccgaaggcgctagggatgggtctcctcaacCCctcctgaaggctggcggcgctcctggTggtaac 2691
I P E G A R D G S P Q P L L K A G G A P G G N
cctaccctttatcaggatcttgttagggTgtcgGCCaccataaccaacactggtaacgTcgccggttat 2760
P T L Y Q D L V R V S A T I T N T G N V A G Y
gaagTccctcaattggTgagTgaccCGcatgttcccttgcgTtgcaatttggttaactcgcttctagtat 2829
E V P Q L Y
gtttcattgggCGgaccgaacgagcctcgggTcgTtctgcgcaagTtCGaccgaatcttccctggctcct 2898
V S L G G P N E P R V V L R K F D R I F L A P
ggggagcaaaaggTttggaccacgactcttaacCGTcgtgatctcgccaattgggatgtggaggctcag 2967
G E Q K V W T T T L N R R D L A N W D V E A Q
gactgggTcatcacaagTaccCCaagaaagTgcacgTcggcagctcctcgcgtaagctgcctctgaga 3036
D W V I T K Y P K K V H V G S S S R K L P L R
gcgcctctgccccgTgtctactag 3060
A P L P R V Y .

```

Fig. 1B

3/6

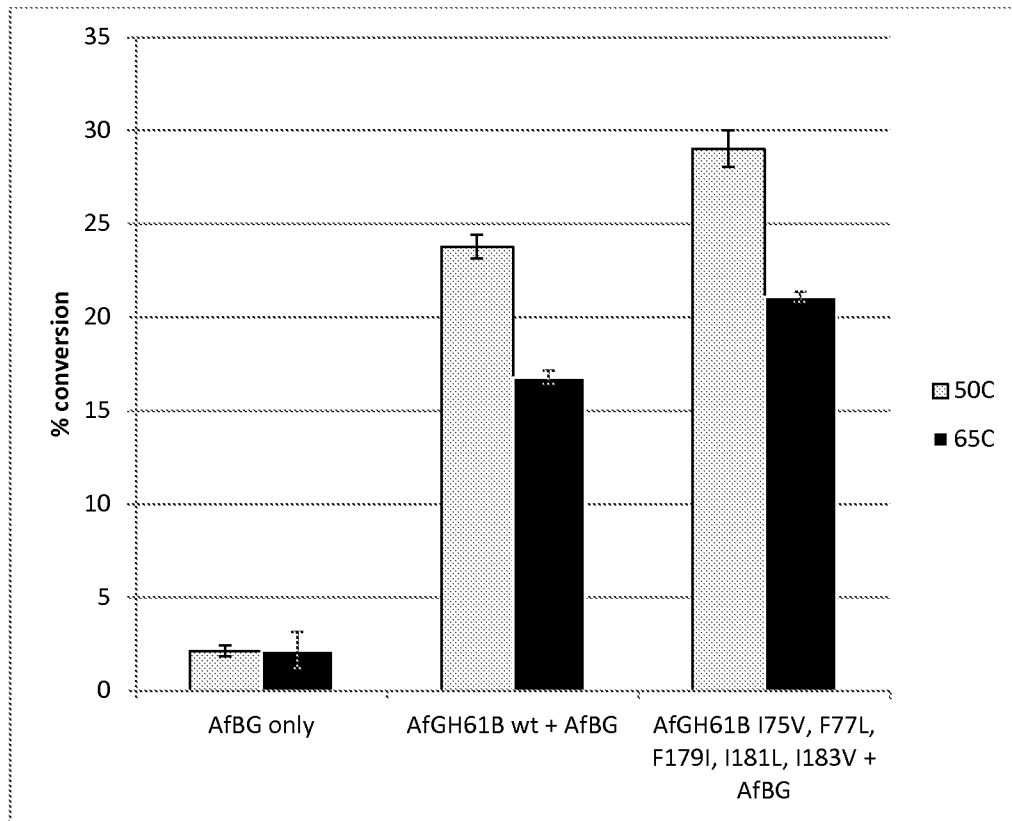


Fig. 2

4/6

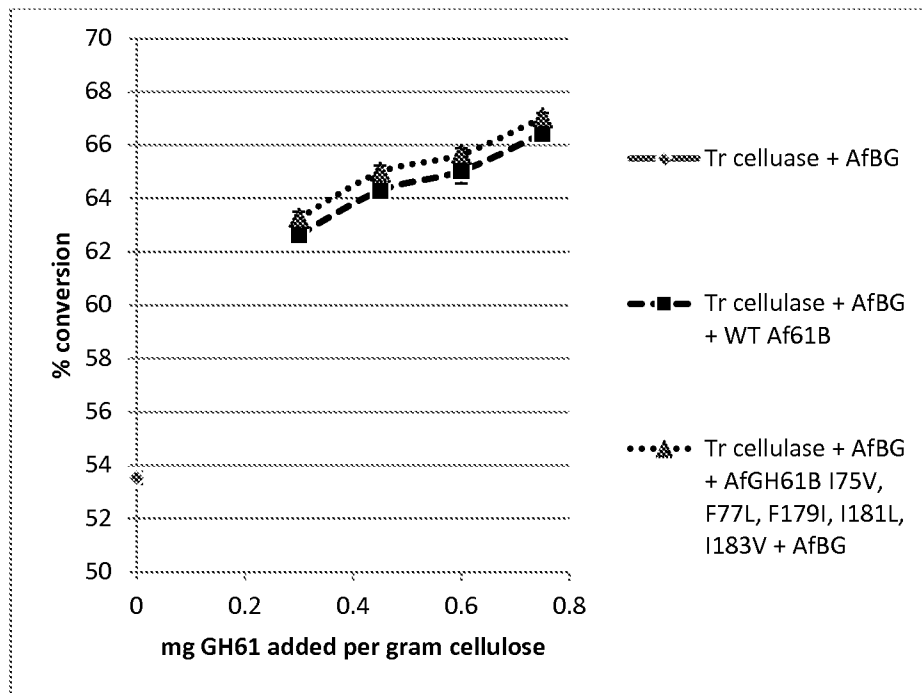


Fig. 3A

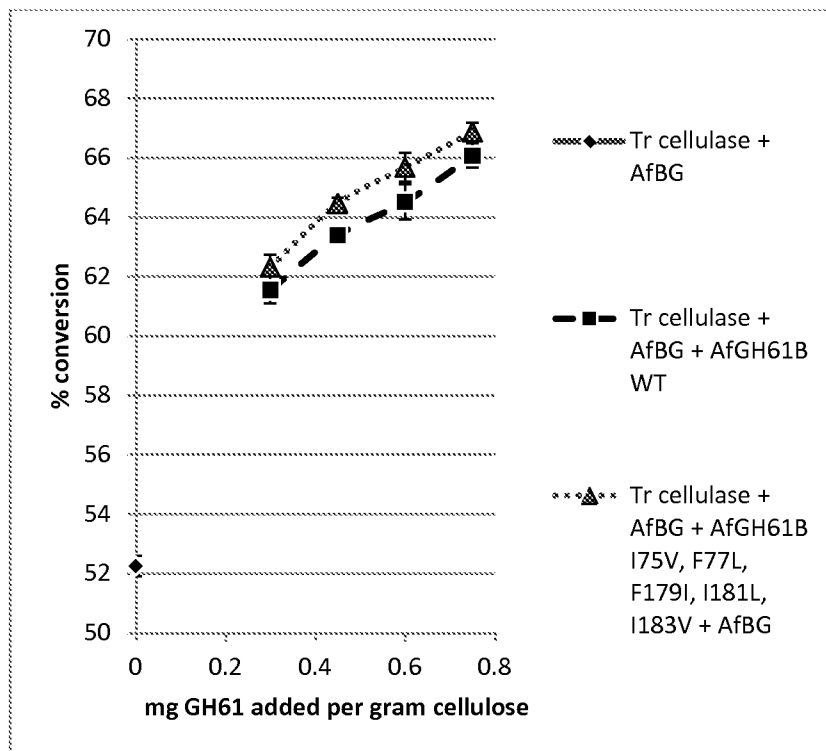


Fig. 3B

5/6

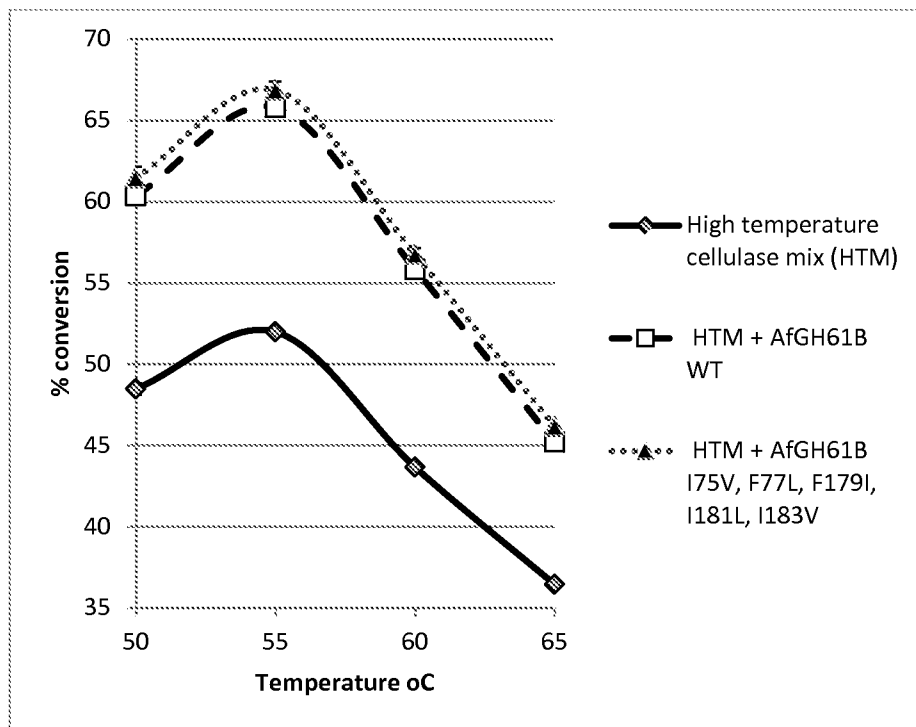


Fig. 4A

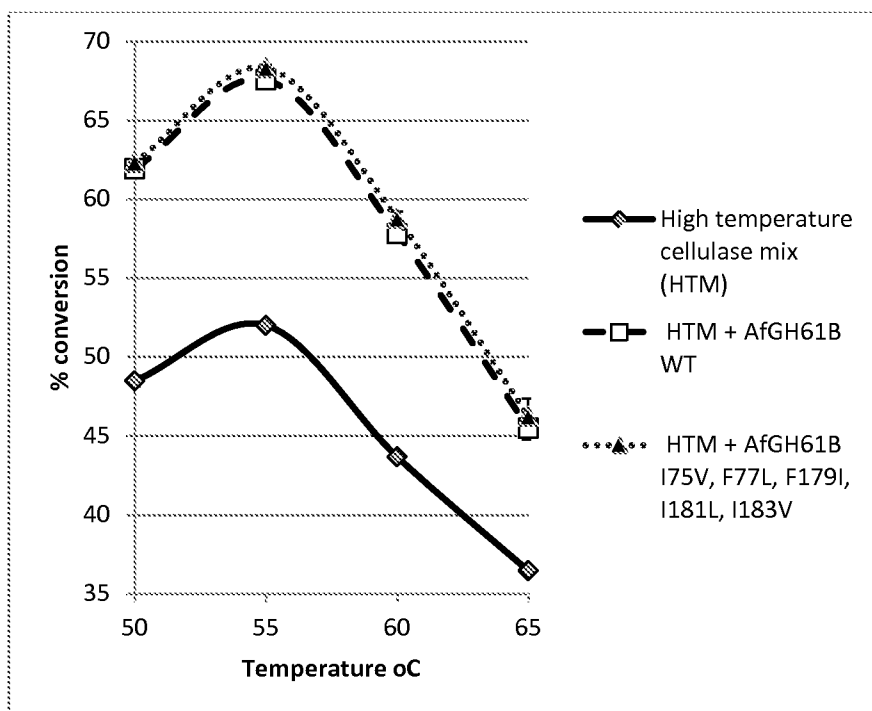


Fig. 4B

6/6

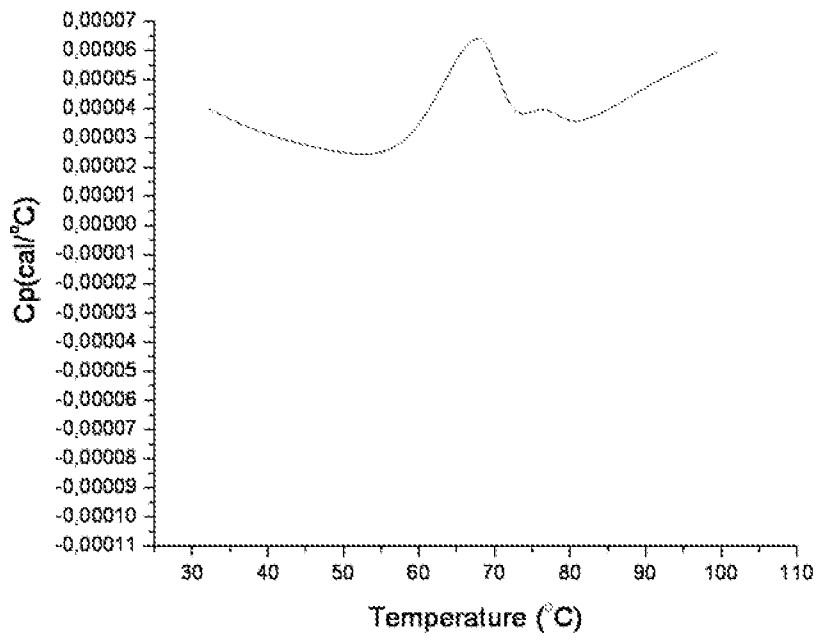


Fig. 5A

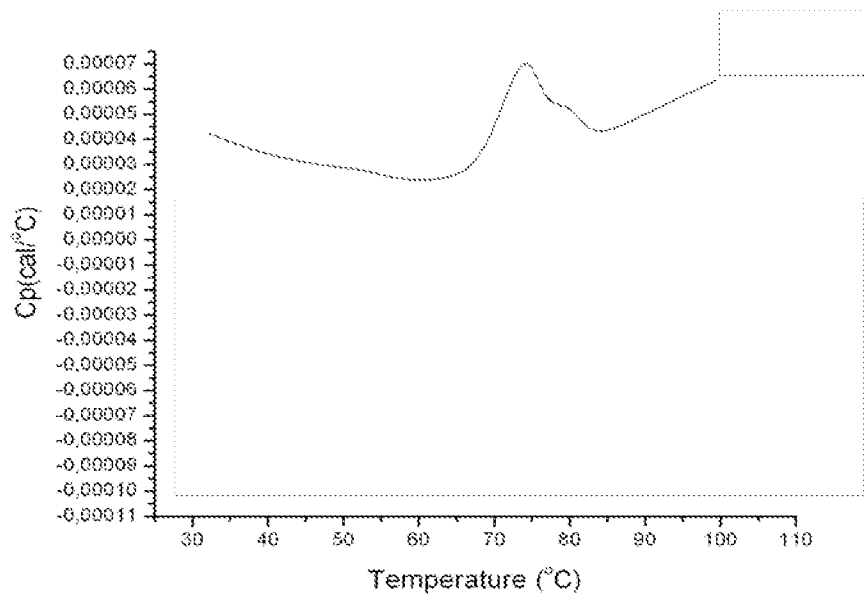


Fig. 5B

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2011/054036

A. CLASSIFICATION OF SUBJECT MATTER
 INV. C12N9/42 C12N15/82 D21C5/00 C12N5/10
 ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
 Minimum documentation searched (classification system followed by classification symbols)
 C12N D21C

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
 EPO-Internal, Sequence Search, BIOSIS, COMPENDEX, EMBASE, FSTA, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE Geneseq [Online] 19 February 2009 (2009-02-19), "T. aurantiacus protein with cellulolytic enhancing activity, SEQ: 12.", XP002667279, retrieved from EBI accession no. GSP:AUP68836 Database accession no. AUP68836 sequence	1-23
X	-& WO 2008/151043 A1 (NOVOZYMES INC [US]; MCFARLAND KEITH [US]; HARRIS PAUL [US]) 11 December 2008 (2008-12-11) the whole document, in particular SEQ ID NO: 12 and claims ----- -/--	1-23

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search 18 January 2012	Date of mailing of the international search report 01/02/2012
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer Bassias, Ioannis

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2011/054036

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE Geneseq [Online]</p> <p>9 July 2009 (2009-07-09), "Thielavia terrestris cellulolytic enhancing activity protein SEQ:64.", XP002667280, retrieved from EBI accession no. GSP:AWW27072 Database accession no. AWW27072 sequence</p>	1-23
X	<p>-& WO 2009/059234 A2 (NOVOZYMES INC [US]; XU FENG [US]) 7 May 2009 (2009-05-07) the whole document, in particular SEQ ID NO: 64 and claims</p>	1-23
X	<p>-----</p> <p>DATABASE UniProt [Online]</p> <p>3 March 2009 (2009-03-03), "SubName: Full=Endoglucanase, putative;", XP002667281, retrieved from EBI accession no. UNIPROT:B8M2G3 Database accession no. B8M2G3 sequence</p>	1-5
X	<p>-----</p> <p>DATABASE UniProt [Online]</p> <p>6 March 2007 (2007-03-06), "SubName: Full=Similarity to hypothetical endoglucanase IV - Trichoderma reesei; Flags: Precursor;", XP002667282, retrieved from EBI accession no. UNIPROT:A2QR94 Database accession no. A2QR94 sequence</p> <p>-& PEL HERMAN J ET AL: "Genome sequencing and analysis of the versatile cell factory Aspergillus niger CBS 513.88", NATURE BIOTECHNOLOGY, NATURE PUBLISHING GROUP, NEW YORK, NY, US, vol. 25, no. 2, 1 February 2007 (2007-02-01), pages 221-231, XP002458027, ISSN: 1087-0156, DOI: 10.1038/NBT1282 the whole document</p> <p>-----</p>	1-5

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No PCT/US2011/054036

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2008151043	A1	11-12-2008	
		AU 2008259937 A1	11-12-2008
		AU 2008259986 A1	11-12-2008
		CA 2689261 A1	11-12-2008
		CA 2689910 A1	11-12-2008
		CN 101809150 A	18-08-2010
		EP 2064323 A1	03-06-2009
		EP 2069492 A2	17-06-2009
		JP 2010528621 A	26-08-2010
		KR 20100020977 A	23-02-2010
		RU 2009149467 A	10-07-2011
		US 2010129860 A1	27-05-2010
		WO 2008151043 A1	11-12-2008
		WO 2008151079 A2	11-12-2008
WO 2009059234	A2	07-05-2009	
		CN 101910406 A	08-12-2010
		EP 2235191 A2	06-10-2010
		US 2009130707 A1	21-05-2009
		WO 2009059234 A2	07-05-2009