Abstract:
The present invention relates to methods for degrading or converting a cellulosic material and for producing substances from the cellulosic material.
METHODS FOR ENHANCING THE DEGRADATION OF CELLULOSIC MATERIAL
WITH CHITIN BINDING PROTEINS

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

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

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 methods for degrading or converting a cellulosic material and for producing substances from the cellulosic material.

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.


There is a need in the art for improving the efficiency of cellulolytic enzyme compositions in the saccharification of cellulosic material.

The present invention provides improved methods for degrading or converting a cellulosic material with an enzyme composition in the presence of a chitin binding protein.

**Summary of the Invention**

The present invention relates to methods for degrading or converting a cellulosic material, comprising: treating the cellulosic material with an enzyme composition in the presence of a chitin binding protein. In one aspect, the methods further comprise recovering the degraded or converted cellulosic material. In another aspect, the cellulosic material is
treated with an enzyme composition in the presence of a chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

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 chitin binding protein; (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. In one aspect, the cellulosic material is saccharified with an enzyme composition in the presence of a chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

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 chitin binding protein. In one aspect, the cellulosic material is saccharified with an enzyme composition in the presence of a chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

In another aspect, the chitin binding protein is selected from the group consisting of:

(a) a chitin binding protein having at least 60% sequence identity to the full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof;

(b) a chitin binding protein encoded by a polynucleotide that hybridizes under at least medium-high stringency conditions with the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; or the full-length complement thereof;

(c) a chitin binding protein encoded by a polynucleotide having at least 60% sequence identity to the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 thereof;

(d) a variant of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24, or the CBM33 thereof, comprising a substitution, deletion, and/or insertion at one or more (e.g., several) positions; and

(e) a fragment of the chitin binding protein of (a), (b), (c), or (d) that has chitin
binding activity.

The present invention also relates to whole broth formulations, cell culture compositions, or enzyme compositions comprising a chitin binding protein or a chitin binding protein and a GH61 polypeptide having cellulytic enhancing activity.

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-napthyl 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 µmol 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 µmol 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 µmol 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 µmol 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.

**Carbohydrate binding module or CBM:** The term "carbohydrate binding module" or "CBM" means a contiguous amino acid sequence within a carbohydrate binding protein with a discreet fold having carbohydrate-binding activity. The term carbohydrate binding module is also referred herein as a chitin binding module.

**CBM33:** The term "CBM33" means a carbohydrate binding module of Family 33, according to the CAZY classification system (Davies and Henrissat, 2002, Biochem. Soc. T30: 291-297 and Bourne and Henrissat, 2001, Curr. Opin. Struct. Biol. 11: 593).

**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 and E.C. 3.2.1.176) that catalyzes the hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, cellooligosaccharides, or any beta-1,4-linked glucose containing polymer, releasing cellobiose from the reducing or non-reducing ends of the chain (Teeri, 1997, Crystalline cellulose degradation: New insight into the function of

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

For purposes of the present invention, cellulolytic enzyme activity is determined by measuring the increase in hydrolysis of a cellulosic material by cellulolytic enzyme(s) under the following conditions: 1-50 mg of cellulolytic enzyme protein/g of cellulose in PCS (or other pretreated cellulosic material) for 3-7 days at a suitable temperature, *e.g.*, 50°C, 55°C, 60°C, or 65°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, 60°C, or 65°C, 72 hours, sugar analysis by AMINEX® HPX-87H column (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

**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, arabinoyxylans,
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, Wiselogel et al., 1995, in Handbook on Bioethanol (Charles E. Wyman,
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.

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.

Chitin: The term "chitin" means any polymer containing beta-(1-4)-N-acetylg glucosamine residues linked in a linear fashion. The term chitin includes without limitation crystalline chitin in the alpha form (chains run anti-parallel), beta form (chains run parallel), gamma form (a mixture of parallel and antiparallel chains), amorphous chitin, colloidal chitin, chitin forms in which part of the N-acetylg glucosamine sugars are deacetylated, and chitosan.

Chitin binding protein or CBP: The term "chitin binding protein" or "CBP" means a protein with binding affinity primarily to chitin (but also various carbohydrates containing N-acetyl-glucosamine or N-acetyl-neuraminic acid subunits). In a preferred aspect, a chitin binding protein comprises or consists of a CBM33. A chitin binding protein may primarily comprise a CBM33 or a CBM33 fused to other carbohydrate binding modules, e.g., CBM2, CBM3, and CBM5, and/or other catalytic proteins. The ability of a chitin binding protein to enhance the hydrolysis of a chitin substrate by, for example, a chitinase, can be determined according to the method described in U.S. Patent Application 20070218046. The ability of a chitin binding protein to enhance the degradation of a cellulosic material by a cellulase composition can be determined according to the Examples described herein. The ability of a chitin binding protein to synergize with a GH61 polypeptide in the degradation of a cellulosic material can be determined according to the Examples described herein.

The chitin binding proteins enhance the hydrolysis of a cellulosic material catalyzed by enzyme having cellulolytic activity by at least 1.01-fold, e.g., at least 1.025-fold, at least 1.05-fold, at least 1.075-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.

The combination of a chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity yield a CBP-GH61 synergistic effect (see Example 9) toward a cellulosic material of at least 1.01, e.g., at least 1.025, at least 1.05, at least 1.075, at least 1.10, at least 1.25, at least 1.5, at least 1.75, at least 2, at least 3, at least 4, at least 5, at least 10, or at least 20.

In one aspect, the chitin binding proteins have at least 20%, e.g., at least 40%, at
least 50%, at least 60%, at least 70%, at least 80%, at least 90%, and at least 100% of the chitin binding activity of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof.

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

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

**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 cellulotic 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 polypeptide 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 polypeptide 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-methoxycinnamamate). 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 \( \mu \text{mole} \) 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. In one aspect, a fragment contains at least 85% of the amino acid residues, e.g., at least 90% of the amino acid residues or at least 95% of the amino acid residues of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof.

**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, Shalom, D. and Shoham, Y. Microbial hemicellulases. *Current Opinion In Microbiology*, 2003, 6(3): 219-228). Hemicellulases are key components in the degradation of plant biomass. Examples of hemicellulases include, but are not limited to, an acetylmannan esterase, an acetylxylan esterase, an arabinanase, an arabinofuranosidase, a coumaric acid esterase, a feruloyl esterase, a galactosidase, a glucuronidase, a glucuronoyl esterase, a mannanase, a mannosidase, a xylanase, and a xylosidase. The substrates of these enzymes, the hemicelluloses, are a heterogeneous group of branched and linear
polysaccharides that are bound via hydrogen bonds to the cellulose microfibrils in the plant cell wall, crosslinking them into a robust network. Hemicelluloses are also covalently attached to lignin, forming together with cellulose a highly complex structure. The variable structure and organization of hemicelluloses require the concerted action of many enzymes for its complete degradation. The catalytic modules of hemicellulases are either glycoside hydrolases (GHs) that hydrolyze glycosidic bonds, or carbohydrate esterases (CEs), which hydrolyze ester linkages of acetate or ferulic acid side groups. These catalytic modules, based on homology of their primary sequence, can be assigned into GH and CE families. Some families, with an overall similar fold, can be further grouped into clans, marked alphabetically (e.g., GH-A). A most informative and updated classification of these and other carbohydrate active enzymes is available in the Carbohydrate-Active Enzymes (CAZy) database. Hemicellulolytic enzyme activities can be measured according to Ghose and Bisaria, 1987, *Pure & Appl. Chem.* 59: 1739-1752, at a suitable temperature, e.g., 50°C, 55°C, 60°C, or 65°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. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

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

**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 chitin binding protein:** The term "mature chitin binding protein" 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 chitin binding protein is amino acids 28 to 206 of SEQ ID NO: 2, amino acids 29 to 194 of SEQ ID NO: 4, amino acids 34 to 201 of SEQ ID NO: 6, amino acids 29 to 220 of SEQ ID NO: 8, amino acids 28 to 478 of SEQ ID NO: 10, amino acids 34 to 285 of SEQ ID NO: 12, amino acids 26 to 199 of SEQ ID NO: 14, amino acids 28 to 197 of SEQ ID NO: 16, amino acids 44 to 491 of SEQ ID NO: 18, amino acids 31 to 201 of SEQ ID NO: 20, amino acids 24 to 487 of SEQ ID NO: 22, or amino acids 22 to 494 of SEQ ID NO: 24 based on the SignalP program (Nielsen et al., 1997, Protein Engineering 10: 1-6) that predicts amino acids 1 to 27 of SEQ ID NO: 2, amino acids 1 to 28 of SEQ ID NO: 4, amino acids 1 to 33 of SEQ ID NO: 6, amino acids 1 to 28 of SEQ ID NO: 8, amino acids 1 to 27 of SEQ ID NO: 10, amino acids 1 to 33 of SEQ ID NO: 12, amino acids 1 to 25 of SEQ ID NO: 14, amino acids 1 to 27 of SEQ ID NO: 16, amino acids 1 to 43 of SEQ ID NO: 18, amino acids 1 to 30 of SEQ ID NO: 20, amino acids 1 to 23 of SEQ ID NO: 22, or amino acids 1 to 21 of SEQ ID NO: 24 are a signal peptide.

**Mature chitin binding protein coding sequence:** The term "mature chitin binding protein coding sequence" means a polynucleotide that encodes a mature chitin binding protein. In one aspect, the mature chitin binding protein coding sequence is nucleotides 82 to 618 of SEQ ID NO: 1, nucleotides 85 to 582 of SEQ ID NO: 3, nucleotides 100 to 603 of SEQ ID NO: 5, nucleotides 85 to 660 of SEQ ID NO: 7, nucleotides 82 to 1434 of SEQ ID NO: 9, nucleotides 100 to 855 of SEQ ID NO: 11, nucleotides 76 to 597 of SEQ ID NO: 13, nucleotides 82 to 591 of SEQ ID NO: 15, nucleotides 130 to 1473 of SEQ ID NO: 17, nucleotides 91 to 603 of SEQ ID NO: 19, nucleotides 70 to 1461 of SEQ ID NO: 21, or nucleotides 64 to 1482 of SEQ ID NO: 23 based on the SignalP program (Nielsen et al., 1997, supra) that predicts nucleotides 1 to 81 of SEQ ID NO: 1, nucleotides 1 to 84 of SEQ ID NO: 3, nucleotides 1 to 99 of SEQ ID NO: 5, nucleotides 1 to 84 of SEQ ID NO: 7, nucleotides 1 to 81 of SEQ ID NO: 9, nucleotides 1 to 99 of SEQ ID NO: 11, nucleotides 1 to 75 of SEQ ID NO: 13, nucleotides 1 to 81 of SEQ ID NO: 15, nucleotides 1 to 129 of SEQ ID NO: 17, nucleotides 1 to 90 of SEQ ID NO: 19, nucleotides 1 to 69 of SEQ ID NO: 21, or nucleotides 1 to 63 of SEQ ID NO: 23 encode a signal peptide.

**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.

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 (e.g., several) 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.

Polypeptide having cellulolytic enhancing activity: The term "polypeptide having cellulolytic enhancing activity" means a GH61 polypeptide that catalyzes the enhancement of the hydrolysis of a cellulosic material by enzyme having cellulolytic activity. For purposes of the present invention, cellulolytic enhancing activity is determined by measuring the increase in reducing sugars or the increase of the total of cellobiose and glucose from the hydrolysis of a cellulosic material by cellulolytic enzyme under the following conditions: 1-50 mg of total protein/g of cellulose in PCS, wherein total protein is comprised of 50-99.5% w/w cellulolytic enzyme protein and 0.5-50% w/w protein of a GH61 polypeptide having cellulolytic enhancing activity for 1-7 days at a suitable temperature, e.g., 50°C, 55°C, 60°C, or 65°C, and pH, e.g., 5.0 or 5.5, compared to a control hydrolysis with equal total protein loading without cellulolytic enhancing activity (1-50 mg of cellulolytic protein/g of cellulose in PCS). In a preferred aspect, a mixture of CELLUCLAST® 1.5L (Novozymes A/S, Bagsvaerd, Denmark) in the presence of 2-3% of total protein weight Aspergillus oryzae beta-glucosidase (recombinantly produced in Aspergillus oryzae according to WO 02/095014) or 2-3% of total protein weight Aspergillus fumigatus beta-glucosidase (recombinantly produced in Aspergillus oryzae as described in WO 2002/095014) of cellulase protein loading is used as the source of the cellulolytic activity.

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 5.0.0 or later. The parameters used are the gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

(Identical Residues x 100)/(Length of Alignment - 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 5.0.0 or later. The parameters used are the gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

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

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 chitin binding protein coding sequence; wherein the subsequence encodes a fragment having chitin binding activity. In one aspect, a subsequence contains at least 85% of the nucleotides, e.g., at least 90% of the nucleotides or at least 95% of the nucleotides of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof.

Variant: The term "variant" means a chitin binding protein comprising an alteration, i.e., a substitution, insertion, and/or deletion at one or more (e.g., several) positions. A
substitution means a 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.

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 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-xidosidases, arabinofuranosidases, alpha-glucuronidases, acetyl/xylan esterases, feruloyl esterases, and alpha-glucuronyl 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-xyllosidase 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-arabinoxytan 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 \( \mu \text{mol} \) 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 \( \mu \text{mol} \) of azurine produced per minute at 37°C, pH 6 from 0.2% AZCL-arabinoxylan as substrate in 200 mM sodium phosphate pH 6 buffer.

**Detailed Description of the Invention**

The present invention relates to methods for degrading or converting a cellulosic material, comprising: treating the cellulosic material with an enzyme composition in the presence of a chitin binding protein. In one aspect, the methods further comprise recovering the degraded or converted cellulosic material. In another aspect, the cellulosic material is treated with an enzyme composition in the presence of a chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

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 chitin binding protein; (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. In one aspect, the cellulosic material is saccharified with an enzyme composition in the presence of a chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

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 chitin binding protein. In one aspect, the cellulosic material is saccharified with an enzyme composition in the presence of a chitin binding protein and a GH61 polypeptide having cellulytic enhancing activity.

The methods of the present invention can be used to saccharify a 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).

10 Chitin Binding Proteins

In an embodiment, the isolated chitin binding proteins have a sequence identity to the full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24, or the CBM33 thereof, 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 chitin binding activity. In one aspect, the chitin binding proteins differ by up to 10 amino acids, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10, from the full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof.

A chitin binding protein in the methods of the present invention preferably comprises or consists of the amino acid sequence of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof; or an allelic variant thereof; or is a fragment thereof retaining chitin binding activity.

In another aspect, the chitin binding protein comprises or consists of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof.

In another aspect, the chitin binding protein comprises or consists of amino acids 28 to 206 of SEQ ID NO: 2, amino acids 29 to 194 of SEQ ID NO: 4, amino acids 34 to 201 of SEQ ID NO: 6, amino acids 29 to 220 of SEQ ID NO: 8, amino acids 28 to 478 of SEQ ID NO: 10, amino acids 34 to 285 of SEQ ID NO: 12, amino acids 26 to 199 of SEQ ID NO: 14, amino acids 28 to 197 of SEQ ID NO: 16, amino acids 44 to 491 of SEQ ID NO: 18, amino
acids 31 to 201 of SEQ ID NO: 20, amino acids 24 to 487 of SEQ ID NO: 22, or amino acids 22 to 494 of SEQ ID NO: 24; or the CBM33 thereof.

In another embodiment, the isolated chitin binding proteins are encoded by polynucleotides that hybridize under very low stringency conditions, low stringency conditions, medium stringency conditions, medium-high stringency conditions, high stringency conditions, or very high stringency conditions with the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23, or the CBM33 coding sequence thereof, or the full-length complement thereof (Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, New York).

The polynucleotide of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23, the mature polypeptide coding sequence thereof, the CBM33 coding sequence thereof, or a subsequence thereof, as well as the chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24, the mature polypeptide thereof, the CBM33 thereof, or a fragment thereof, may be used to design nucleic acid probes to identify and clone DNA encoding chitin binding proteins from strains of different genera or species according to methods well known in the art. In particular, such probes can be used for hybridization with the genomic DNA or cDNA of a cell of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Such probes can be considerably shorter than the entire sequence, but should be at least 15, e.g., at least 25, at least 35, or at least 70 nucleotides in length. Preferably, the nucleic acid probe is at least 100 nucleotides in length, e.g., at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, at least 500 nucleotides, at least 600 nucleotides, at least 700 nucleotides, at least 800 nucleotides, or at least 900 nucleotides in length. Both DNA and RNA probes can be used. The probes are typically labeled for detecting the corresponding gene (for example, with $^{32}$P, $^3$H, $^{35}$S, biotin, or avidin). Such probes are encompassed by the present invention.

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 chitin binding protein. Genomic or other DNA from such other strains may be separated by agarose or polyacrylamide gel electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to and immobilized on nitrocellulose or other suitable carrier material. In order to identify a clone or DNA that is homologous with SEQ ID NO: 1,
the mature chitin binding protein coding sequence thereof; the CBM33 coding sequence thereof; or a subsequence thereof; the carrier material is preferably used in a Southern blot.

For purposes of the present invention, hybridization indicates that the polynucleotide hybridizes to a labeled nucleic acid probe corresponding to SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; the mature chitin binding protein coding sequence thereof; the CBM33 coding sequence thereof; a full-length complement thereof; or a subsequence thereof; under very low to very high stringency conditions. Molecules to which the nucleic acid probe hybridizes under these conditions can be detected using, for example, X-ray film or any other detection means known in the art.

In one aspect, the nucleic acid probe is a polynucleotide that encodes the chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; the mature chitin binding protein thereof; the CBM33 thereof; or a fragment thereof.

In another aspect, the nucleic acid probe is SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; the mature chitin binding protein coding sequence thereof; or the CBM33 coding sequence thereof.

In another embodiment, the isolated chitin binding proteins are encoded by polynucleotides having a sequence identity to the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23, or the CBM33 coding sequence thereof, 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%.

In another embodiment, the isolated chitin binding proteins are variants of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24, or the CBM33 thereof, comprising a substitution, deletion, and/or insertion at one or more (e.g., several) positions. Preferably, amino acid changes are 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.

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.
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.

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 activity to identify amino acid residues that are critical to the activity
of the molecule. See also, Hilton et al., 1996, J. Biol. Chem. 271: 4699-4708. The active site
of the enzyme or other biological interaction can also be determined by physical analysis of
structure, as determined by such techniques as nuclear magnetic resonance,
crystallography, electron diffraction, or photoaffinity labeling, in conjunction with mutation of
putative contact site amino acids. See, for example, de Vos et al., 1992, Science 255: 306-
59-64. The identity of essential amino acids can also be inferred from an alignment with a
related polypeptide.

Single or multiple amino acid substitutions, deletions, and/or insertions can be made
and tested using known methods of mutagenesis, recombination, and/or shuffling, followed
by a relevant screening procedure, such as those disclosed by Reidhaar-Olson and Sauer,
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,

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 ai, 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.

In an embodiment, the number of amino acid substitutions, deletions, and/or insertions introduced into the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24, or the CBM33 thereof, is up to 10, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

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


Additional examples of chitin binding proteins that may be used in the methods of the present invention are listed below with their accession numbers, which are incorporated herein by reference. It is understood herein that each of the chitin binding proteins below are included in each of the embodiments above.

- 21 -
Alferomonas sp. 0-7 (GENBANK AB063629, GENPEPT BAB79619.1)
Burkholderia mallei ATCC 23344 (GENBANK CP000001, GENPEPT AAU45854.1)
Burkholderia mallei ATCC 23344 (GENBANK CP000010, GENPEPT AAU48386.1)
Burkholderia pseudo mallei 1710b (GENBANK CP000124, GENPEPT ABA49030.1)
Burkholderia pseudo mallei 1710b (GENBANK CP000125, GENPEPT ABA53645.1)
Burkholderia pseudo mallei K96243 (GENBANK BX571965, GENPEPT CAH37353.1)
Burkholderia pseudo mallei K96243 (GENBANK BX571966, GENPEPT CAH37950.1)
Burkholderia sp. 383 (GENBANK CP000150, GENPEPT ABB05775.1)
Burkholderia thailandensis E264; ATCC 700388 (GENBANK CP000085, GENPEPT ABC34637.1)
Burkholderia thailandensis E264; ATCC 700388 (GENBANK CP000086, GENPEPT ABC38514.1)
Caldibacillus cellulovorans (GENBANK AF163837, GENPEPT AAF22274.1)
Chromobacterium violaceum ATCC 12472 (GENBANK AE01691, GENPEPT AAQ58230.1, GENBANK NC_005085, GENPEPT NP 900224.1)
Chromobacterium violaceum ATCC 12472 (GENBANK AE01691, GENPEPT AAQ58229.1, GENBANK NC_005085, GENPEPT NP 900223.1)
Chromobacterium violaceum ATCC 12472 (GENBANK AE016919, GENPEPT AAQ60987.1, GENBANK NC_005085, GENPEPT NP 902993.1)
Chromobacterium violaceum ATCC 12472 (GENBANK AE016921, GENPEPT AAQ60262.1, GENBANK NC_005085, GENPEPT NP 902262.1)
Chromobacterium violaceum ATCC 12472 (GENBANK AE016922, GENPEPT AAQ61150.1, GENBANK NC_005085, GENPEPT NP 903159.1)
Enterococcus faecalis V583 (GENBANK AE016948, GENPEPT AA080225.1, GENBANK NC_004668, GENPEPT NP 814154.1)
Enterococcus faecium (GENPEPT AAQ43729.1)
Francisella tularensis subsp. holarctica LVS (GENBANK AM233362, GENPEPT CAJ79847.1)
Francisella tularensis subsp. tularensis Schu 4 (GENBANK AJ749949, GENPEPT CAG45449.1)
Hahella chejuensis KCTC 2396 (GENBANK CP000015, GENPEPT ABC27701.1)
Hahella chejuensis KCTC 2396 (GENBANK CP000015, GENPEPT ABC30692.1)
Lactobacillus plantarum WCFS 1 (GENBANK AL935256, GENPEPT CAD64126.1, GENBANK NC_004567, GENPEPT NP 785278.1)
Lactobacillus sakei subsp. sakei 23K (GENBANK CR936503, GENPEPT
CA1 5531 0.1 )

Lactococcus lactis subsp. lactis IL 1403 (GENBANK AE006425, GENPEPT AAK06049.1, GENBANK NC_002662, GENPEPT NP 268108.1)

Legionella pneumophila Paris (GENBANK CR628336, GENPEPT CAH1 1404.1)

Listeria innocua (GENBANK AL596173, GENPEPT CAC97838.1, GENBANK NC_003212, GENPEPT NP_471941.1)

Listeria monocytogenes EGD-e (GENBANK AL591983, GENPEPT CAD00545.1, GENBANK NC_003210, GENPEPT NP_465990.1)

Listeria monocytogenes 4b F2365 (GENBANK AE017330, GENPEPT AAT05205.1)

Oceanobacillus iheyensis HTE831 (GENBANK AP004595, GENPEPT BAC12766.1, GENBANK NC_004193, GENPEPT NP_691731.1)

Photobacterium profundum SS9 CR378676 CAG22185.1 (GENBANK CR378676, GENPEPT CAG22185.1)

Photorhabdus luminescens subsp. laumondii TT01 (GENBANK BX571866, GENPEPT CAE14645.1, GENBANK NC_005126, GENPEPT NP 929598.1)

Proteus mirabilis (GENPEPT AAR43285.1)

Pseudoalteromonas sp. S9 (GENBANK AF007895, GENPEPT AAC79666.1)

Pseudomonas aeruginosa PA01 (GENBANK AE004520, GENPEPT AAG04241.1, GENBANK NC_002516, GENPEPT NP 249543.1)

Pseudomonas aeruginosa PA025 (GENBANK AF1 96565, GENPEPT AAF12807.1)

Pseudomonas fluorescens Pf-5 (GENBANK CR000076, GENPEPT AAY91365.1)

Pseudomonas fluorescens PfO-1 (GENBANK CP000094, GENPEPT ABA75307.1)

Pseudomonas syringae pv. syringae B728a (GENBANK CP000075, GENPEPT AAY37892.1)

Pseudomonas syringae pv. tomato DC3000 (GENBANK AE016866, GENPEPT AA056470.1, GENBANK NC_004578, GENPEPT NP 792775.1)

Rickettsia felis URRWXCal2 (GENBANK CP000053, GENPEPT AAY61559.1)

Saccharophagus degradans 2-40 (GENBANK BK001045, GENPEPT DAA01337.1)

Salinivibrio costicola 5SM-1 (GENBANK GYA207003, GENPEPT AAP42509.1)

Serratia marcescens 2170 (GENBANK AB01 5998, GENPEPT BAA31569.1)

Serratia marcescens BJL200 (GENBANK AY665558, GENPEPT AUA88202.1)

Serratia marcescens KCTC2172 (GENBANK L38484, GENPEPT AAC37123.1)

Shewanella oneidensis MR-1 (GENBANK AE015551, GENPEPT AAN54144.1, GENBANK NC_004347, GENPEPT NP 716699.1)

Sodalis glossinidius 'morsitans' (GENBANK AP008232, GENPEPT BAE74790.1)

Streptomyces avermitilis MA-4680 (GENBANK AP005047, GENPEPT BAC74271.1, GENBANK NC_003155, GENPEPT NP 827736.1)
Streptomyces avermitilis MA-4680 (GENBANK AP005029, GENPEPT BAC69879.1, GENBANK NC_003155, GENPEPT NP 823344.1)

Streptomyces avermitilis MA-4680 (GENBANK AP005042, GENPEPT BAC72935.1, GENBANK NC_003155, GENPEPT NP 826400.1)

Streptomyces avermitilis MA-4680 (GENBANK AP005030, GENPEPT BAC69965.1, GENBANK NC_003155, GENPEPT NP 823430.1)

Streptomyces coelicolor A3(2) (GENBANK AL359215, GENPEPT CAB94648.1, GENBANK NC_003888, GENPEPT NP 631281.1)

Streptomyces coelicolor A3(2) (GENBANK AL031155, GENPEPT CAA20076.1, GENBANK NC_003888, GENPEPT NP 626400.1)

Streptomyces coelicolor A3(2) (GENBANK AL136058, GENPEPT CAB65563.1, GENBANK NC_003888, GENPEPT NP 627062.1)

Streptomyces coelicolor A3(2) (GENBANK AL132973, GENPEPT CAB61160.1, GENBANK NC_003888, GENPEPT NP 624952.1)

Streptomyces coelicolor A3(2) (GENBANK AB017013, GENPEPT BAA75647.1, GENBANK AL121719, GENPEPT CAB57190.1, GENBANK NC_003888, GENPEPT NP 624799.1)

Streptomyces coelicolor A3(2) (GENBANK AL096849, GENPEPT CAA55284.1, GENBANK NC_003888, GENPEPT NP 626007.1)

Streptomyces coelicolor A3(2) (GENBANK AL133210, GENPEPT CAB61600.1, GENBANK NC_003888, GENPEPT NP 625478.1)

Streptomyces griseus (GENBANK AB023785, GENPEPT BAA86267.1)

Streptomyces halstedii (GENBANK U51222, GENPEPT AAC45430.1)

Streptomyces olivaceoviridis ATCC 11238 (GENBANK X78535, GENPEPT CAA55284.1)

Streptomyces retculi (GENBANK Y14315, GENPEPT CAA74695.1)

Streptomyces thermoviolaceus OPC-520 (GENBANK AB 11 0078, GENPEPT BAD01591.1)

Streptomyces viridosporus (GENBANK AF126376, GENPEPT AAD27623.1)

Thermobifida fusca YX (GENBANK CP000088, GENPEPT AAZ55700.1)

Thermobifida fusca YX (GENBANK CP000088, GENPEPT AAZ55306.1)

Vibrio cholerae N16961 (GENBANK AE004355, GENPEPT AAF96053.1, GENBANK NC 002506, GENPEPT NP 232540.1)

Vibrio cholerae N16961 (GENBANK AE004409, GENPEPT AAF96709.1, GENBANK NC 002506, GENPEPT NP 233197.1)

Vibrio fischeri ES1 14 (GENBANK CP000021, GENPEPT AAW87213.1)

Vibrio fischeri ES1 14 (GENBANK CP000021, GENPEPT AAW87083.1)
Vibrio parahaemolyticus RIMD 2210633 (GENBANK AP005084, GENPEPT BAC61435.1, GENBANK NC 004605, GENPEPT NP 799602.1)

Vibrio parahaemolyticus RIMD 2210633 (GENBANK AP005089, GENPEPT BAC62941.1, GENBANK NC 004605, GENPEPT NP 801108.1)

Vibrio vulnificus CMCP6 (GENBANK AE016812, GENPEPT AA008152.1, GENBANK NC 004460, GENPEPT NP 763162.1)

Vibrio vulnificus CMCP6 (GENBANK AE016808, GENPEPT AA007021.1, GENBANK NC 004460, GENPEPT NP 762031.1)

Vibrio vulnificus YJ016 (GENBANK AP005344, GENPEPT BAC96112.1, GENBANK NC 005140, GENPEPT NP 936142.1)

Vibrio vulnificus YJ016 (GENBANK AP005346, GENPEPT BAC96577.1, GENBANK NC 005140, GENPEPT NP 936607.1)

Yersinia enterocolitica (type 0:8) WA-314 (GENBANK AJ344214, GENPEPT CAC83040.2)

Yersinia pestis biovar Medievalis 91001 (GENBANK AE017129, GENPEPT AAS60972.1, GENBANK NC 005810, GENPEPT NP 992095.1)

Yersinia pestis C092 (GENBANK AJ414156, GENPEPT CAC92462.1, GENBANK NC 003143, GENPEPT NP 406699.1)

Yersinia pestis KIM (GENBANK AE013699, GENPEPT AAM84543.1, GENBANK NC 004088, GENPEPT NP 668292.1)

Yersinia pseudotuberculosis IP 32953 (GENBANK BX936398, GENPEPT CAH22604.1)

Yersinia pseudotuberculosis IP 32953 (GENBANK BX936398, GENPEPT CAH20139.1)

Agrotis segetum nucleopolyhedrovirus (GENBANK DQ123841, GENPEPT AAZ38192.1)

Autographa californica nucleopolyhedrovirus (GENBANK D00583, GENPEPT BAA00461.1, GENBANK L22858, GENPEPT AAA66694.1, GENBANK NC 001962, GENPEPT NP 047468.1)

Bombyx mori nuclear polyhedrosis virus (GENBANK U55071, GENPEPT AAB47606.1, GENBANK NC 001962, GENPEPT NP 047468.1, GENBANK L33180, GENPEPT AAC63737.1)

Choristoneura biennis entomopoxvirus (GENBANK M34140, GENPEPT AAA42887.1)

Choristoneura furniferana defective nucleopolyhedrovirus (GENBANK AY327402, GENPEPT AAQ91667.1, GENBANK NC 005137, GENPEPT NP 932669.1)

Choristoneura furniferana nuclear polyhedrosis virus (GENBANK U26734,
Chrysodeixis chalcites nucleopolyhedrovirus (GENBANK AY864330, GENPEPT AAY83998.1)

Epiphyas postvittana nucleopolyhedrovirus (GENBANK AY043265, GENPEPT AAK85621.1, GENBANK NC 003083, GENPEPT NP 203226.1)

Helicoverpa armigera single nucleocapsid polyhedrovirus (GENBANK AF266696, GENPEPT AAK57880.1, GENBANK AF303045, GENPEPT AAK96305.1, GENBANK NC 003094, GENPEPT NP 203613.1)

Helicoverpa zea nucleopolyhedrovirus (GENBANK AF334030, GENPEPT AAL56204.1, GENBANK NC 003349, GENPEPT NP 542682.1)

Helicoverpa armigera nucleopolyhedrovirus G4 (GENBANK AF271059, GENPEPT AAG53801.1, GENBANK NC 002654, GENPEPT NP 075127.1)

Heliothis armigera entomopoxvirus (GENBANK L08077, GENPEPT AAA92858.1)

Hyphantria cunea nucleopolyhedrovirus (GENBANK AP009046, GENPEPT BAE72375.1)

Leucania separata nuclear polyhedrosis virus (GENBANK AB009614, GENPEPT BAA24259.1)

Lymantria dispar nucleopolyhedrovirus (GENBANK U38895, GENPEPT AAB07702.1, GENBANK AF081810, GENPEPT AAC70254.1, GENBANK NC 001973, GENPEPT NP 047705.1)

Mamestra brassicae nucleopolyhedrovirus (GENBANK AF108960, GENPEPT AAD45231.1)

Mamestra configurata nucleopolyhedrovirus A (GENBANK U59461, GENPEPT AAM09 145.1, GENBANK AF539999, GENPEPT AAQ1 1056.1)

Mamestra configurata nucleopolyhedrovirus B (GENBANK AY126275, GENPEPT AAM9501 9.1, GENBANK NC 0041 17, GENPEPT NP 689207.1)

Orgyia pseudotsugata nuclear polyhedrosis virus (GENBANK U75930, GENPEPT AAC59068.1, GENBANK D13306, GENPEPT BAA02566.1, GENBANK NC 001875, GENPEPT NP 046225.1)

Pseudaletia separata entomopoxvirus (GENBANK DS0590, GENPEPT BAA09138.1)

Spodoptera exigua nucleopolyhedrovirus (GENBANK AF1 69823, GENPEPT AAF33555.1, GENBANK NC 002169, GENPEPT NP 037785.1)

Spodoptera frugiperda MNPV (GENBANK AY250076, GENPEPT AAP79107.10)

Spodoptera litura nucleopolyhedrovirus G2 (GENBANK AF325155, GENPEPT AAL01718.1, GENBANK NC 003102, GENPEPT NP 258300.1)

Trichoplusia ni single nucleopolyhedrovirus (GENBANK DQ017380, GENPEPT AAZ67435.1)
Unidentified entomopoxvirus (GENBANK X77616, GENPEPT CAA54706.1)

*Xestia c-nigrum* granulovirus (GENBANK AAF05221, GENPEPT CAA54706.1, GENBANK NC 002331, GENPEPT NP 059255.1)

5 Sources of Chitin Binding Proteins

A chitin binding protein may be obtained from organisms and 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 chitin binding protein 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 chitin binding protein obtained from a given source is secreted extracellularly.

The chitin binding protein may be a bacterial chitin binding protein. For example, the chitin binding protein may be a Gram-positive bacterial polypeptide such as a *Bacillus*, *Caldibacillus*, *Chromobacterium*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Listeria*, *Lysinibacillus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, *Streptomyces*, or *Thermobifida* polypeptide, or a Gram-negative bacterial polypeptide such as an *Alferomonas*, *Burkholderia*, *Caldibacillus*, *Campylobacter*, *E. coli*, *Flavobacterium*, *Francisella*, *Fusobacterium*, *Hahella*, *Helicobacter*, *Ilyobacter*, *Legionella*, *Neisseria*, *Photobacterium*, *Proteus*, *Pseudoalteromonas*, *Pseudomonas*, *Photorhabdus*, *Rickettsia*, *Saccharophagus*, *Salinivibrio*, *Salmonella*, *Serratia*, *Shewanella*, *Sodalis*, *Ureaplasma*, *Vibrio*, or *Yersinia* polypeptide.

In one aspect, the chitin binding protein is a *Bacillus* *alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus anthracis*, *Bacillus brevis*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus cytotoxicus*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus mycoides*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, *Bacillus thuringiensis*, *Bacillus weihenstephanensis*, or *Lysinibacillus sphaericus* polypeptide.

In another aspect, the chitin binding protein is a *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, or *Streptococcus equi* subsp. *Zooepidemicus* polypeptide.

In another aspect, the chitin binding protein is a *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, or *Streptomyces lividans* polypeptide.

The chitin binding protein may be a fungal chitin binding protein. For example, the chitin binding protein may be a yeast polypeptide such as a *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* polypeptide; or a filamentous fungal polypeptide such as an *Acremonium*, *Agaricus*, *Alternaria*, *Aspergillus*, *Aurantiporus*,...

In another aspect, the chitin binding protein is a Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norbensis, or Saccharomyces oviformis polypeptide.


The chitin binding protein may also be an insect, mammalian, plant, or virus chitin binding protein.

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

The chitin binding protein may be identified and obtained from other sources including microorganisms isolated from nature (e.g., soil, composts, water, etc.) using the above-mentioned probes. Techniques for isolating microorganisms from natural habitats are well known in the art. A polynucleotide encoding the chitin binding protein 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 chitin binding protein has been detected with the probe(s), the polynucleotide can be isolated or cloned by utilizing techniques that are well known to those of ordinary skill in the art (see, e.g., Sambrook et al., 1989, supra).

**Polynucleotides**

Polynucleotides encoding chitin binding proteins can be isolated and utilized to practice the methods of the present invention, as described herein.

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

Modification of a polynucleotide encoding a chitin binding protein may be necessary for synthesizing chitin binding proteins substantially similar to the chitin binding protein. The term “substantially similar” to the chitin binding protein refers to non-naturally occurring forms of the polypeptide. These polypeptides may differ in some engineered way from the polypeptide isolated from its native source, e.g., variants that differ in specific activity, thermostability, pH optimum, or the like. The variants may be constructed on the basis of the polynucleotide presented as the mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID
NO: 23, or the CBM33 coding sequence thereof, or a subsequence thereof, and/or by introduction of nucleotide substitutions that do not result in a change in the amino acid sequence of the polypeptide, but which correspond to the codon usage of the host organism intended for production of the enzyme, or by introduction of nucleotide substitutions that may give rise to a different amino acid sequence. For a general description of nucleotide substitution, see, e.g., Ford et al., 1991, Protein Expression and Purification 2: 95-107.

**Nucleic Acid Constructs**

A polynucleotide encoding a chitin binding protein or an enzyme of interest may be operably linked to one or more (e.g., several) control sequences that direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences.

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

The control sequence may be a promoter, a polynucleotide that is recognized by a host cell for expression of a polynucleotide encoding the chitin binding protein. The promoter contains transcriptional control sequences that mediate the expression of the chitin binding protein. 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.


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, which is recognized by a host cell to terminate transcription. The terminator is operably linked to the 3'-terminus of the polynucleotide encoding the chitin binding protein. Any terminator that is functional in the host cell may be used in the present invention.

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

Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* acetamidase, *Aspergillus nidulans* anthranilate synthase, *Aspergillus*

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

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

Examples of suitable mRNA stabilizer regions are obtained from a *Bacillus thuringiensis crylIIA* 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 is operably linked to the 5'-terminus of the polynucleotide encoding the chitin binding protein. 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 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 polynucleotide and, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence that is functional in the host cell may be used.

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

The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a chitin binding protein and directs the chitin binding protein into the cell's secretory pathway. The 5'-end of the coding sequence of the polynucleotide may inherently contain a signal peptide coding sequence naturally linked in translation reading frame with the segment of the coding sequence that encodes the chitin binding protein. 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 chitin binding protein. However, any signal peptide coding sequence that directs the expressed chitin binding protein 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 chitin binding protein. 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 chitin binding protein 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 at the N-terminus of a chitin binding protein, the propeptide sequence is positioned next to the N-terminus of a
chitin binding protein 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
chitin binding protein 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-amyrase 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 chitin binding protein
would be operably linked to the regulatory sequence.

Expression Vectors

A polynucleotide encoding a chitin binding protein or an enzyme of interest and
various nucleic acids and control sequences described herein 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 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, 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 (phosphoribosylaminimidazole-succinocarboxamide synthase), adeB (phosphoribosyl-aminimidazole synthase), amdS (acetamidase), argB (ornithine carbamoyltransferase), bar (phosphinotricin acetyltransferase), hph (hygromycin phosphotransferase), niaD (nitrate reductase), pyrG (orotidine-5'-phosphate decarboxylase), sC (sulfate adenyltransferase), and trpC (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are *Aspergillus nidulans* or *Aspergillus oryzae* amdS and pyrG genes and a *Streptomyces hygroscopicus* bar gene. Preferred for use in a *Trichoderma* cell are adeA, adeB, amdS, hph, and pyrG genes.

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

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 pAMβI 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 ANSI (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 may be inserted into a host cell to increase production of a chitin binding protein. 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 are well known to one skilled in the art (see, e.g., Sambrook et al., 1989, supra).

Host Cells

Recombinant host cells comprising a polynucleotide encoding a chitin binding protein or an enzyme of interest operably linked to one or more (e.g., several) control sequences that direct the production of a chitin binding protein can be advantageously used in the recombinant production of the chitin binding protein. 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 chitin binding protein and its source.

The host cell may be any cell useful in the recombinant production of a chitin binding protein, 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 lavidus* cells.


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

The host cell may be a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota as 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).

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).

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 kluveri, Saccharomyces norbensis, Saccharomyces oviformis*, or *Yarrowia lipolytica* cell.

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

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


**Methods of Production**

A chitin binding protein can be produced using methods comprising: (a) cultivating a cell, which in its wild-type form produces the chitin binding protein, under conditions conducive for production of the chitin binding protein; and optionally (b) recovering the chitin binding protein.

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

The chitin binding protein may be detected using methods known in the art that are specific for the chitin binding proteins. These detection methods include, but are not limited to, use of specific antibodies, adsorption by chitin, enhancement of chitinase reaction on chitin, or specific activity on chitin. For example, an enzyme assay based on oxidative chitin degradation may be used to determine the amount or activity of the chitin binding protein (Vanje-Kolstad et al., 2010, Science 330: 219).

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

The chitin binding protein 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 chitin binding proteins.

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

An enzyme of interest can also be produced, recovered, and/or purified by the methods described above.

Fermentation Broth Formulations or Cell Compositions

The present invention also relates to a fermentation broth formulation or a cell composition comprising a chitin binding protein. In one aspect, the fermentation broth
formulation or a cell composition comprises a chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity. The fermentation broth product further comprises additional ingredients used in the fermentation process, such as, for example, cells (including, the host cells containing the gene encoding the polypeptide of the present invention which are used to produce the polypeptide of interest), cell debris, biomass, fermentation media and/or fermentation products. In some embodiments, the composition is a cell-killed whole broth containing organic acid(s), killed cells and/or cell debris, and culture medium.

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 or bacterial 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 (e.g., several) additional enzymes selected from the group consisting of a cellulase, a GH61 polypeptide having cellulolytic enhancing activity, a hemicellulase, an esterase, an expansin,
a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

The cell-killed whole broth or composition may contain the unfractionated contents of the fermentation materials derived at the end of the fermentation. Typically, the cell-killed whole broth or composition contains the spent culture medium and cell debris present after the microbial cells (e.g., filamentous fungal or bacterial cells) are grown to saturation, incubated under carbon-limiting conditions to allow protein synthesis. In some embodiments, the cell-killed whole broth or composition contains the spent cell culture medium, extracellular enzymes, and killed filamentous fungal or bacterial 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, 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.

**Enzyme Compositions**

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

The compositions may comprise a chitin binding protein as the major component, e.g., a mono-component composition. Alternatively, the compositions may comprise multiple enzymatic activities, such as one or more (e.g., several) additional enzymes selected from the group consisting of a cellulase, a GH61 polypeptide having cellulolytic enhancing activity, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin. In one aspect, the composition comprises a chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

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.

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

The processing of a cellulosic material according to the methods of 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. The production of a desired fermentation product from the cellulosic material typically involves pretreatment, enzymatic hydrolysis (saccharification), and fermentation.

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

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, Flavio Faria de Moraes, Gisella Maria

Additional reactor types include: fluidized bed, upflow blanket, immobilized, and extruder type reactors for hydrolysis and/or fermentation.


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 CO$_2$, supercritical H$_2$O, 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 S0₂ (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 2006/1 10891, WO 2006/1 10899, WO 2006/1 10900, and WO 2006/1 10901 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.


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.

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 chitin binding protein or a chitin binding protein and a GH61 polypeptide. The enzymes 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(s), i.e., optimal for the enzyme(s). The hydrolysis can be carried out as a fed batch or continuous process where the cellulosic material is fed gradually to, for example, an enzyme containing hydrolysis solution.

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

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 GH61 polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises an endoglucanase and a GH61 polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises a cellobiohydrolase and a GH61 polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises a beta-glucosidase and a GH61 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 GH61 polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises an endoglucanase, a beta-glucosidase, and a GH61 polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises a cellobiohydrolase, a beta-glucosidase, and a GH61 polypeptide having cellulolytic enhancing activity. In another aspect, the enzyme composition comprises a GH61 polypeptide having cellulolytic enhancing activity or a GH61 polypeptide is added to the chitin binding protein, which synergizes with the chitin binding protein in the degradation or conversion of a cellulosic material.
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 \( \text{H}_2\text{O}_2 \) 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 polylol, and/or lactic acid or another organic acid according to established processes.

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

In one aspect, an effective amount of cellulolytic or hemicellulolytic enzyme to the cellulosic material is about 0.5 to about 50 mg, e.g., about 0.5 to about 40 mg, about 0.5 to about 25 mg, about 0.75 to about 20 mg, about 0.75 to about 15 mg, about 1.0 to about 10 mg, about 1.5 to about 10 mg, or about 2.5 to about 10 mg per g of the cellulosic material.

In another aspect, an effective amount of a chitin binding protein to the cellulosic material is about 0.01 to about 50 mg, e.g., about 0.01 to about 40 mg, about 0.01 to about 30 mg, about 0.01 to about 20 mg, about 0.01 to about 10 mg, about 0.01 to about 5 mg, about 0.025 to about 1.5 mg, about 0.05 to about 1.25 mg, about 0.075 to about 1.25 mg, about 0.1 to about 1.25 mg, about 0.15 to about 1.25 mg, or about 0.25 to about 1.0 mg per g of the cellulosic material.

In another aspect, an effective amount of a GH61 polypeptide having cellulolytic enhancing activity to the cellulosic material is about 0.01 to about 50 mg, e.g., about 0.01 to about 40 mg, about 0.01 to about 30 mg, about 0.01 to about 20 mg, about 0.01 to about 10 mg, about 0.01 to about 5 mg, about 0.025 to about 1.5 mg, about 0.05 to about 1.25 mg, about 0.075 to about 1.25 mg, about 0.1 to about 1.25 mg, about 0.15 to about 1.25 mg, or about 0.25 to about 1.0 mg per g of the cellulosic material.

In another aspect, an effective amount of a chitin binding protein to cellulolytic enzyme is about 0.005 to about 1.0 g, e.g., about 0.01 to about 1.0 g, about 0.05 to about 0.75 g, about 0.05 to about 0.5 g, about 0.1 to about 0.5 g, about 0.1 to about 0.25 g, or about 0.05 to about 0.2 g per g of cellulolytic enzyme.

In another aspect, an effective amount of a GH61 polypeptide having cellulolytic enhancing activity to cellulolytic enzyme is about 0.005 to about 1.0 g, e.g., about 0.01 to about 1.0 g, about 0.05 to about 0.75 g, about 0.05 to about 0.5 g, about 0.1 to about 0.5 g, about 0.1 to about 0.25 g, or about 0.05 to about 0.2 g per g of cellulolytic enzyme.

In another aspect, an effective amount of a chitin binding protein to a GH61 polypeptide having cellulolytic enhancing activity is in a ratio (wt/wt) of about 0.01 to about 100, e.g., about 0.1 to about 10, about 0.2 to about 5, about 0.5 to about 2, or about 1 g per g of GH61 polypeptide having cellulolytic enhancing activity.
The polypeptides having cellulolytic enzyme activity or hemicellulolytic enzyme activity, as well as other proteins/polypeptides, e.g., GH61 polypeptide having cellulolytic enhancing activity, useful in the degradation of the cellulosic material (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, Thermotoga, 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 steathermophilus, 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,

In one aspect, the polypeptide is a Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasi, Saccharomyces kluveri, Saccharomyces norbensis, or Saccharomyces oviformis 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 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), CELLIC® CTec3 (Novozymes A/S), CELLUCLAST™ (Novozymes A/S), NOVOZYM™ 188 (Novozymes A/S), CELLUZYME™ (Novozymes A/S), CEREFLO™ (Novozymes A/S), and ULTRAFLO™ (Novozymes A/S), ACCELERASE™ (Genencor Int.), LAMINEX™ (Genencor Int.), SPEZYME™ CP (Genencor Int.), FILTRASE® NL (DSM); METHAPLUS® S/L 100 (DSM), ROHAMENT™ 7069 W (Rohm GmbH), FIBREZYME® LDI (Dyadic International, Inc.), FIBREZYME® LBR (Dyadic International, Inc.), or VISCOSTAR® 150L (Dyadic International, Inc.). The cellulase enzymes are added in amounts effective from about 0.001 to about 5.0 wt % of solids, e.g., about 0.025 to about 4.0 wt % of solids or about 0.005 to about 2.0 wt % of solids.

In the enzyme compositions of the present invention, any GH61 polypeptide having cellulolytic enhancing activity can be used.

In a first aspect, isolated polypeptides having cellulolytic enhancing activity, comprise the following motifs:

\[ [ILMV]-P-X(4,5)-G-X-Y-[ILMV]-X-R-X-[EQ]-X(4)-[HNQ] \] (SEQ ID NO: 25 or SEQ ID NO: 26) and \[ [FW]-[TF]-K-[AIV], \]

wherein X is any amino acid, X(4,5) is any amino acid at 4 or 5 contiguous positions, and X(4) is any amino acid at 4 contiguous positions.

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

\[ H-X(1,2)-G-P-X(3)-[YW]-[AILMV] \] (SEQ ID NO: 27 or SEQ ID NO: 28),
\[ [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV] \] (SEQ ID NO: 29), or
\[ H-X(1,2)-G-P-X(3)-[YW]-[AILMV] \] (SEQ ID NO: 30 or SEQ ID NO: 31) and \[ [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV] \] (SEQ ID NO: 32),

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

In a preferred embodiment, the isolated GH61 polypeptide having cellulolytic enhancing activity further comprises \[ H-X(1,2)-G-P-X(3)-[YW]-[AILMV] \] (SEQ ID NO: 27 or SEQ ID NO: 28). In another preferred embodiment, the isolated GH61 polypeptide having cellulolytic enhancing activity further comprises \[ [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV] \] (SEQ ID NO: 29). In another preferred embodiment, the isolated GH61 polypeptide having cellulolytic enhancing activity further comprises \[ H-X(1,2)-G-P-X(3)-[YW]-[AILMV] \] (SEQ ID NO: 30 or SEQ ID NO: 31) and \[ [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV] \] (SEQ ID NO: 32).
In a second aspect, isolated polypeptides having cellulolytic enhancing activity, comprise the following motif:

\[ [\text{ILMV}]-\text{P-X(4,5)-G-X-Y-}[\text{ILMV}]-\text{X-R-X-}[\text{EQ}]-\text{X(3)-A-[HNO]} \]  
(SEQ ID NO: 33 or SEQ ID NO: 34),

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

In a third aspect, the GH61 polypeptide having cellulolytic enhancing activity comprises or consists of an amino acid sequence 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%, or at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% sequence identity to the mature polypeptide of 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, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, SEQ ID NO: 70, SEQ ID NO: 72, SEQ ID NO: 74, SEQ ID NO: 76, SEQ ID NO: 78, SEQ ID NO: 80, SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 112, SEQ ID NO: 114, SEQ ID NO: 116, SEQ ID NO: 118, SEQ ID NO: 120, SEQ ID NO: 122, or SEQ ID NO: 124.

In a fourth aspect, the GH61 polypeptide having cellulolytic enhancing activity is encoded by a polynucleotide that hybridizes under at least very low stringency conditions, e.g., at least low stringency conditions, at least medium stringency conditions, at least medium-high stringency conditions, at least high stringency conditions, or at least very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, SEQ ID NO: 71, SEQ ID NO: 73, SEQ ID NO: 75, SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, SEQ ID NO: 111, SEQ ID NO: 113, SEQ ID NO: 115, SEQ ID NO: 117, SEQ ID NO: 119, SEQ ID NO: 121, or SEQ ID NO: 123, (ii) the genomic DNA sequence of the mature polypeptide coding sequence of SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 49, SEQ ID NO:
115, SEQ ID NO: 117, or SEQ ID NO: 119, or the cDNA sequence of the mature polypeptide coding sequence of SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 47, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, SEQ ID NO: 71, SEQ ID NO: 73, SEQ ID NO: 75, SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, SEQ ID NO: 111, SEQ ID NO: 113, SEQ ID NO: 121, or SEQ ID NO: 123, or (iii) a full-length complement of (i) or (ii) (J. Sambrook, E.F. Fritsch, and T. Maniatis, 1989, supra).

In a fifth aspect, the GH61 polypeptide having cellulolytic enhancing activity is encoded by a polynucleotide comprising or consisting of a nucleotide sequence 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: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, SEQ ID NO: 71, SEQ ID NO: 73, SEQ ID NO: 75, SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, SEQ ID NO: 111, SEQ ID NO: 113, SEQ ID NO: 115, SEQ ID NO: 117, SEQ ID NO: 119, SEQ ID NO: 121, or SEQ ID NO: 123.

In a sixth aspect, the GH61 polypeptide having cellulolytic enhancing activity is a variant of the mature polypeptide of 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, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, SEQ ID NO: 70, SEQ ID NO: 72, SEQ ID NO: 74, SEQ ID NO: 76, SEQ ID NO: 78, SEQ ID NO: 80, SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 112, SEQ ID NO: 114, SEQ ID NO: 116, SEQ ID NO: 118, SEQ ID NO: 120, SEQ ID NO: 122, or SEQ ID NO: 124 comprising a substitution, deletion, and/or insertion at one or more (e.g., several) positions.

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

Examples of conservative substitutions are within the group of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine and asparagine), hydrophobic amino acids (leucine, isoleucine and valine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine, serine, threonine and methionine). Amino acid substitutions that do not generally alter specific activity are known in the art and are described, for example, by H. Neurath and R.L. Hill, 1979, In, The Proteins, Academic Press, New York. The most commonly occurring exchanges 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, LeuA/αl, 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.

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

Single or multiple amino acid substitutions, deletions, and/or insertions can be made and tested using known methods of mutagenesis, recombination, and/or shuffling, followed by a relevant screening procedure, such as those disclosed by Reidhaar-Olson and Sauer, 1988, Science 241: 53-57; Bowie and Sauer, 1989, Proc. Natl. Acad. Sci. USA 86: 2152-
Mutagenesis/shuffling methods can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides expressed by host cells (Ness et al., 1999, *Nature Biotechnology* 17: 893-896). Mutagenized DNA molecules that encode active polypeptides can be recovered from the host cells and rapidly sequenced using standard methods in the art. These methods allow the rapid determination of the importance of individual amino acid residues in a polypeptide.

The total number of amino acid substitutions, deletions and/or insertions of the mature polypeptide of 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, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, SEQ ID NO: 70, SEQ ID NO: 72, SEQ ID NO: 74, SEQ ID NO: 76, SEQ ID NO: 78, SEQ ID NO: 80, SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 112, SEQ ID NO: 114, SEQ ID NO: 116, SEQ ID NO: 118, SEQ ID NO: 120, SEQ ID NO: 122, or SEQ ID NO: 124 is up to 10, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

In an embodiment, the mature polypeptide comprises or consists of amino acids 20 to 326 of SEQ ID NO: 36, amino acids 18 to 239 of SEQ ID NO: 38, amino acids 20 to 258 of SEQ ID NO: 40, amino acids 19 to 226 of SEQ ID NO: 42, amino acids 20 to 304 of SEQ ID NO: 44, amino acids 16 to 317 of SEQ ID NO: 46, amino acids 22 to 249 of SEQ ID NO: 48, amino acids 20 to 249 of SEQ ID NO: 50, amino acids 18 to 232 of SEQ ID NO: 52, amino acids 16 to 235 of SEQ ID NO: 54, amino acids 19 to 323 of SEQ ID NO: 56, amino acids 16 to 310 of SEQ ID NO: 58, amino acids 20 to 246 of SEQ ID NO: 60, amino acids 22 to 354 of SEQ ID NO: 62, amino acids 22 to 250 of SEQ ID NO: 64, amino acids 22 to 322 of SEQ ID NO: 66, amino acids 24 to 444 of SEQ ID NO: 68, amino acids 26 to 253 of SEQ ID NO: 70, amino acids 18 to 246 of SEQ ID NO: 72, amino acids 20 to 334 of SEQ ID NO: 74, amino acids 18 to 227 of SEQ ID NO: 76, amino acids 20 to 223 of SEQ ID NO: 78, amino acids 22 to 368 of SEQ ID NO: 80, amino acids 25 to 330 of SEQ ID NO: 82, amino acids 17 to 236 of SEQ ID NO: 84, amino acids 19 to 250 of SEQ ID NO: 86, amino acids 23 to 478 of SEQ ID NO: 88, amino acids 17 to 230 of SEQ ID NO: 90, amino acids 20 to 257 of SEQ ID NO: 92, amino acids 23 to 251 of SEQ ID NO: 94, amino acids 19 to 349 of SEQ ID NO: 96, amino acids 24 to 436 of SEQ ID NO: 98, amino acids 21 to 344 of SEQ ID NO: 100, amino acids...

In another embodiment, the mature polypeptide coding sequence comprises or consists of is nucleotides 388 to 1332 of SEQ ID NO: 35 or the cDNA sequence thereof, nucleotides 98 to 821 of SEQ ID NO: 37 or the cDNA sequence thereof, nucleotides 126 to 978 of SEQ ID NO: 39 or the cDNA sequence thereof, nucleotides 55 to 678 of SEQ ID NO: 41 or the genomic DNA sequence thereof, nucleotides 58 to 912 of SEQ ID NO: 43 or the genomic DNA sequence thereof, nucleotides 46 to 951 of SEQ ID NO: 45 or the genomic DNA sequence thereof, nucleotides 64 to 796 of SEQ ID NO: 47 or the cDNA sequence thereof, nucleotides 77 to 766 of SEQ ID NO: 49 or the genomic DNA sequence thereof, nucleotides 52 to 921 of SEQ ID NO: 51 or the cDNA sequence thereof, nucleotides 46 to 851 of SEQ ID NO: 53 or the cDNA sequence thereof, nucleotides 55 to 1239 of SEQ ID NO: 55 or the cDNA sequence thereof, nucleotides 46 to 1250 of SEQ ID NO: 57 or the cDNA sequence thereof, nucleotides 58 to 811 of SEQ ID NO: 59 or the cDNA sequence thereof, nucleotides 64 to 1112 of SEQ ID NO: 61 or the cDNA sequence thereof, nucleotides 64 to 859 of SEQ ID NO: 63 or the cDNA sequence thereof, nucleotides 64 to 1018 of SEQ ID NO: 65 or the cDNA sequence thereof, nucleotides 70 to 1483 of SEQ ID NO: 67 or the cDNA sequence thereof, nucleotides 76 to 832 of SEQ ID NO: 69 or the cDNA sequence thereof, nucleotides 52 to 875 of SEQ ID NO: 71 or the cDNA sequence thereof, nucleotides 58 to 1250 of SEQ ID NO: 73 or the cDNA sequence thereof, nucleotides 52 to 795 of SEQ ID NO: 75 or the cDNA sequence thereof, nucleotides 58 to 974 of SEQ ID NO: 77 or the cDNA sequence thereof, nucleotides 64 to 1104 of SEQ ID NO: 79 or the cDNA sequence thereof, nucleotides 73 to 990 of SEQ ID NO: 81 or the cDNA sequence thereof, nucleotides 49 to 1218 of SEQ ID NO: 83 or the cDNA sequence thereof, nucleotides 55 to 930 of SEQ ID NO: 85 or the cDNA sequence thereof, nucleotides 67 to 1581 of SEQ ID NO: 87 or the cDNA sequence thereof, nucleotides 49 to 865 of SEQ ID NO: 89 or the cDNA sequence thereof, nucleotides 58 to 1065 of SEQ ID NO: 91 or the cDNA sequence thereof, nucleotides 67 to 868 of SEQ ID NO: 93 or the cDNA sequence thereof, nucleotides 55 to 1099 of SEQ ID NO: 95 or the cDNA sequence thereof, nucleotides 70 to 1483 of SEQ ID NO: 97 or the cDNA sequence thereof, nucleotides 61 to 1032 of SEQ ID NO: 99 or the cDNA sequence thereof, nucleotides 76 to 1200 of SEQ ID NO: 101 or the cDNA sequence thereof, nucleotides 61 to 1167 of SEQ ID NO: 103 or the cDNA sequence thereof, nucleotides 64 to 1218 of SEQ ID NO: 105 or the cDNA sequence thereof, nucleotides 58 to
1281 of SEQ ID NO: 107 or the cDNA sequence thereof, nucleotides 52 to 801 of SEQ ID NO: 109 or the cDNA sequence thereof, nucleotides 61 to 819 of SEQ ID NO: 111 or the cDNA sequence thereof, nucleotides 61 to 966 of SEQ ID NO: 113 or the cDNA sequence thereof, nucleotides 52 to 702 of SEQ ID NO: 115 or the genomic DNA sequence thereof, nucleotides 70 to 699 of SEQ ID NO: 117 or the genomic DNA sequence thereof, nucleotides 49 to 711 of SEQ ID NO: 119 or the genomic DNA sequence thereof, nucleotides 76 to 1452 of SEQ ID NO: 121 or the cDNA sequence thereof, or nucleotides 64 to 1018 of SEQ ID NO: 123 or the cDNA sequence thereof.

In the methods of the present invention, a GH61 polypeptide having cellulolytic enhancing activity of the present invention 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 GH61 polypeptide having cellulolytic enhancing activity is used in the presence of a dioxy compound, a bicyclic compound, a heterocyclic compound, a nitrogen-containing compound, a quinone compound, a sulfur-containing compound, or a liquor obtained from a pretreated cellulosic material such as pretreated corn stover (PCS).

The dioxy compound may include any suitable compound containing two or more oxygen atoms. In some aspects, the dioxy compounds contain a substituted aryl moiety as described herein. The dioxy compounds may comprise one or more (e.g., several) hydroxyl and/or hydroxyl derivatives, but also include substituted aryl moieties lacking hydroxyl and hydroxyl derivatives. Non-limiting examples of the dioxy compounds include pyrocatechol or catechol; caffeic acid; 3,4-dihydroxybenzoic acid; 4-tert-butyl-5-methoxy-1,2-benzenediol; pyrogallol; gallic acid; methyl-3,4,5-trihydroxybenzoate; 2,3,4-trihydroxybenzophenone; 2,6-dimethoxyphenol; sinapinic acid; 3,5-dihydroxybenzoic acid; 4-chloro-1,2-benzenediol; 4-nitro-1,2-benzenediol; tannic acid; ethyl gallate; methyl glycolate; dihydroxyfumaric acid; 2-butyne-1,4-diol; (croconic acid; 1,3-propanediol; tartaric acid; 2,4-pentanediol; 3-ethoxy-1,2-propanediol; 2,4,4′-trihydroxybenzophenone; cis-2-butene-1,4-diol; 3,4-dihydroxy-3-cyclobutene-1,2-dione; dihydroxyacetone; acrolein acetel; 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 flavlyium 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, dihydrothiено-pyrazolyl, thianaphthenyl, carbazolyl, benzimidazolyl, benzothienyl, benzofuranyl, indolyl, quinolyl, benzoquinolyl, benzothiazolyl, benzoazolyl, benzimidazolyl, isoquinolyl, isoindolyl, acridinyl, benzoisoxazolyl, dimethylhydantoin, pyrazinyl, tetrahydrofuranyl, pyrrolinyl, pyrrolidinyl, morpholinyl, indolyl, diazepinyl, azipinyl, thiepinyl, piperidinyl, and oxepinyl. In another aspect, the optionally substituted heterocycloalkyl moiety or optionally substituted heteroaryl moiety is an optionally substituted furanyl. Non-limiting examples of the heterocyclic compounds include

(1,2-dihydroxyethyl)-3,4-dihydroxyfuran-2(5H)-one; 4-hydroxy-5-methyl-3-furanone; 5-hydroxy-2(5H)-furanone; [1,2-dihydroxyethyl]furan-2,3,4(5H)-trione; a-hydroxy-V-butylrolactone; ribonic γ-lactone; aldohexuronicaldohexuronic acid γ-lactone; gluconic acid δ-lactone; 4-hydroxycoumarin; dihydrobenzofuran; 5-(hydroxymethyl)furfural; furon; 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 (e.g., several) nitrogen atoms. In one aspect, the nitrogen-containing compound comprises an amine, imine, hydroxylamine, or nitrooxide 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 Q_{10}; 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 (e.g., several) 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 10⁶ to about 10, e.g., about 10⁻⁶ to about 7.5, about 10⁻⁶ to about 5, about 10⁻⁵ to about 2.5, about 10⁻⁴ to about 1, about 10⁻³ to about 10⁻¹, about 10⁻² to about 10⁻¹, about 10⁻² to about 1, about 10⁻¹ to about 10⁻², or about 10⁻¹ to about 10⁻⁰. In another aspect, an effective amount of such a compound described above is about 0.1 µM to about 1 M, e.g., about 0.5 µM to about 0.75 M, about 0.75 µM to about 0.5 M, about 1 µM to about 0.25 M, about 1 µM to about 0.1 M, about 5 µM to about 50 mM, about 10 µM to about 25 mM, about 50 µM to about 25 mM, about 10 µM to about 10 mM, about 5 µM 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 g, about 10⁻⁶ to about 2.5 g, about 10⁻⁶ to about 1 g, about 10⁻⁵ to about 1 g, about 10⁻⁵ to about 10⁻⁴ g, about 10⁻⁴ to about 10⁻³ g, about 10⁻³ to about 10⁻² g, or about 10⁻² g per g of cellulose.

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

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: 126); Trichoderma reesei endoglucanase II (Saloheimo, et al., 1988, Gene 63:1 1-22); Trichoderma reesei Cel5A endoglucanase II (GENBANK™ accession no. M19373; SEQ ID NO: 128); Trichoderma reesei endoglucanase III (Okada et al., 1988, Appl. Environ. Microbiol. 64: 555-563; GENBANK™ accession no. AB003694; SEQ ID NO: 130); Trichoderma reesei endoglucanase V (Saloheimo et al., 1994, Molecular Microbiology 13: 219-228; GENBANK™ accession no. Z33381; SEQ ID NO: 132); 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 carotovara endoglucanase (Saarilaiti 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: 134); Myceliophthora thermophila CBS 117.65 endoglucanase (SEQ ID NO: 136); basidiomycete CBS 495.95 endoglucanase (SEQ ID NO: 138); basidiomycete CBS 494.95 endoglucanase (SEQ ID NO: 140); Thielavia terrestris NRRL 8126 CEL6B endoglucanase (SEQ ID NO: 142); Thielavia terrestris NRRL 8126 CEL6C endoglucanase (SEQ ID NO: 144); Thielavia terrestris NRRL 8126 CEL7C endoglucanase (SEQ ID NO: 146); Thielavia terrestris NRRL 8126 CEL7E endoglucanase (SEQ ID NO: 148); Thielavia terrestris NRRL 8126 CEL7F endoglucanase (SEQ ID NO: 150); Cladorrhinum foecundissimum ATCC 62373 CEL7A endoglucanase (SEQ ID NO: 152); and Trichoderma reesei strain No. VTT-D-80133 endoglucanase (GENBANK™ accession no. M15665; SEQ ID NO: 154). The endoglucanases of SEQ ID NO: 126, SEQ ID NO: 128, SEQ ID NO: 130, SEQ ID NO: 132, SEQ ID NO: 134, SEQ ID NO: 136, SEQ ID NO: 138, SEQ ID NO: 140, SEQ ID NO: 142, SEQ ID NO: 144, SEQ ID NO: 146, SEQ ID NO: 148, SEQ ID NO: 150, SEQ ID NO: 152, and SEQ ID NO: 154 described above are encoded by the mature polypeptide coding sequence of SEQ ID NO: 125, SEQ ID NO: 127, SEQ ID NO: 129, SEQ ID NO: 131, SEQ ID NO: 133, SEQ ID NO: 135, SEQ ID NO: 137, SEQ ID NO: 139, SEQ ID NO: 141, SEQ ID NO: 143, SEQ ID NO: 145, SEQ ID NO: 147, SEQ ID NO: 149, SEQ ID NO: 151, and SEQ ID NO: 153, respectively.

Examples of cellobiohydrolases useful in the present invention include, but are not
limited to, *Trichoderma reesei* cellobiohydrolase I (SEQ ID NO: 156); *Trichoderma reesei* cellobiohydrolase II (SEQ ID NO: 158); *Humicola insolens* cellobiohydrolase I (SEQ ID NO: 160); *Myceliophthora thermophila* cellobiohydrolase II (SEQ ID NO: 162 and SEQ ID NO: 164); *Thielavia terrestris* cellobiohydrolase II (CEL6A) (SEQ ID NO: 166); *Chaetomium thermophilum* cellobiohydrolase I (SEQ ID NO: 168); *Chaetomium thermophilum* cellobiohydrolase II (SEQ ID NO: 170), *Aspergillus fumigatus* cellobiohydrolase I (SEQ ID NO: 172), and *Aspergillus fumigatus* cellobiohydrolase II (SEQ ID NO: 174). The cellobiohydrolases of SEQ ID NO: 156, SEQ ID NO: 158, SEQ ID NO: 160, SEQ ID NO: 162, SEQ ID NO: 164, SEQ ID NO: 166, SEQ ID NO: 168, SEQ ID NO: 170, SEQ ID NO: 172, and SEQ ID NO: 174 described above are encoded by the mature polypeptide coding sequence of SEQ ID NO: 157, SEQ ID NO: 159, SEQ ID NO: 161, SEQ ID NO: 163, SEQ ID NO: 165, SEQ ID NO: 167, SEQ ID NO: 169, SEQ ID NO: 171, and SEQ ID NO: 173, respectively.

Examples of beta-glucosidases useful in the present invention include, but are not limited to, *Aspergillus oryzae* beta-glucosidase (SEQ ID NO: 176); *Aspergillus fumigatus* beta-glucosidase (SEQ ID NO: 178); *Penicillium brasiliannum* IBT 20888 beta-glucosidase (SEQ ID NO: 180); *Aspergillus niger* beta-glucosidase (SEQ ID NO: 182); and *Aspergillus aculeatus* beta-glucosidase (SEQ ID NO: 184). The beta-glucosidases of SEQ ID NO: 176, SEQ ID NO: 178, SEQ ID NO: 180, SEQ ID NO: 182, and SEQ ID NO: 184 described above are encoded by the mature polypeptide coding sequence of SEQ ID NO: 175, SEQ ID NO: 177, SEQ ID NO: 179, SEQ ID NO: 181, and SEQ ID NO: 183, 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 brasiliannum* 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.

The beta-glucosidase may be a fusion protein. In one aspect, the beta-glucosidase is the *Aspergillus oryzae* beta-glucosidase variant BG fusion protein of SEQ ID NO: 186 or the *Aspergillus oryzae* beta-glucosidase fusion protein of SEQ ID NO: 188 obtained according to WO 2008/057637. The beta-glucosidase fusion proteins of SEQ ID NO: 186 and SEQ ID NO: 188 are encoded by SEQ ID NO: 185 and SEQ ID NO: 187, respectively.


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), CELLC™ HTec (Novozymes A/S), CELLC™ 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, Aspergillus aculeatus xylanase (GeneSeqP:AAR63790; WO 94/21785), Aspergillus fumigatus xylanases (WO 2006/078256; SEQ ID NO: 190), and Thielavia terrestris NRRL 8126 xylanases (WO 2009/079210).

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: 192), Talaromyces emersonii beta-xylosidase (SwissProt accession number Q8X212), and Neurospora crassa beta-xylosidase (SwissProt accession number Q7SOW4).

Examples of acetylxyan esterases useful in the methods of the present invention include, but are not limited to, Hypocrea jecorina acetylxyan esterase (WO 2005/001036), Neurospora crassa acetylxyan esterase (UniProt accession number q7s259), Thielavia terrestris NRRL 8126 acetylxyan esterase (WO 2009/042846), Chaetomium globosum acetylxyan esterase (Uniprot accession number Q2GWX4), Chaetomium gracile acetylxyan esterase (GeneSeqP accession number AAB82124), Phaeosphaeria nodorum acetylxyan esterase (Uniprot accession number QOUH1J), and Humicola insolens DSM 1800 acetylxyan esterase (WO 2009/073709).

Examples of ferulic acid esterases useful in the methods of the present invention include, but are not limited to, Humicola insolens DSM 1800 feruloyl esterase (WO
2009/076122), Neurospora crassa feruloyl esterase (UniProt accession number Q9HGR3), and Neosartorya fischeri feruloyl esterase (UniProt Accession number A1D9T4).

Examples of arabinofuranosidases useful in the methods of the present invention include, but are not limited to, Humicola insolens DSM 1800 arabinofuranosidase (WO 2009/073383) and Aspergillus niger arabinofuranosidase (GeneSeqP accession number AAR94170).

Examples of alpha-glucuronidases useful in the methods of the present invention include, but are not limited to, Aspergillus clavatus alpha-glucuronidase (UniProt accession number alcc12), Trichoderma reesei alpha-glucuronidase (UniProt accession number Q99024), Talaromyces emersonii alpha-glucuronidase (UniProt accession number Q8X211), Aspergillus niger alpha-glucuronidase (UniProt accession number Q96WX9), Aspergillus terreus alpha-glucuronidase (SwissProt accession number Q0CJP9), and Aspergillus fumigatus alpha-glucuronidase (SwissProt accession number Q4WW45).

In a preferred embodiment, the enzyme composition is a high temperature composition, i.e., a composition that is able to hydrolyze a cellulosic material in the range of about 54°C to about 65°C. In another preferred embodiment, the enzyme composition is a high temperature composition, i.e., a composition that is able to hydrolyze a cellulosic material at a temperature of about 54°C, about 55°C, about 56°C, about 57°C, about 58°C, about 59°C, about 60°C, about 61°C, about 62°C, about 63°C, about 64°C, or about 65°C. In another preferred embodiment, the enzyme composition is a high temperature composition, i.e., a composition that is able to hydrolyze a cellulosic material at a temperature of at least 54°C, at least 55°C, at least 56°C, at least 57°C, at least 58°C, at least 59°C, at least 60°C, at least 61°C, at least 62°C, at least 63°C, at least 64°C, or at least 65°C.

In another preferred embodiment, the enzyme composition is a high temperature composition as disclosed in PCT/US2010/055723 (WO 2011/057140), which is incorporated herein in its entirety by reference.

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).

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

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 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.

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).

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

In a preferred aspect, the yeast is a Bretannomyces. In a more preferred aspect, the yeast is a Bretannomyces clausenii. In another preferred aspect, the yeast is a Candida. In another more preferred aspect, the yeast is Candida sonorensis. In another more preferred aspect, the yeast is Candida boidinii. In another more preferred aspect, the yeast is Candida blankii. In another more preferred aspect, the yeast is Candida brassicae. In another more preferred aspect, the yeast is Candida ditdensiae. In another more preferred aspect, the yeast is Candida entomophila. In another more preferred aspect, the yeast is Candida pseudotropicalis. In another more preferred aspect, the yeast is Candida scehatae. 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 Klyveromyces. In another more preferred aspect, the yeast is Klyveromyces fragilis. In another more preferred aspect, the yeast is Klyveromyces
marxianus. In another more preferred aspect, the yeast is Kluyveromyces thermotolerans. In another preferred aspect, the yeast is a Pachysolen. In another more preferred aspect, the yeast is Pachysolen tannophilus. In another preferred aspect, the yeast is a Pichia. In another more preferred aspect, the yeast is a Pichia stipitis. In another preferred aspect, the yeast is a Saccharomyces spp. In another more preferred aspect, the yeast is Saccharomyces cerevisiae. In another more preferred aspect, the yeast is Saccharomyces distaticus. In another more preferred aspect, the yeast is Saccharomyces uvarum.

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 Clostridium. In another more preferred aspect, the bacterium is Clostridium acetobutylicum. In another 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.


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 x 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.

A fermentation stimulator can be used in combination with any of the processes described herein to further improve the fermentation process, and in particular, the performance of the fermenting microorganism, such as, rate enhancement and ethanol yield. A "fermentation stimulator" refers to stimulators for growth of the fermenting microorganisms, in particular, yeast. Preferred fermentation stimulators for growth include vitamins and minerals. Examples of vitamins include multivitamins, biotin, pantothenate, nicotinic acid,
meso-inositol, thiamine, pyridoxine, para-aminobenzoic acid, folic acid, riboflavin, and Vitamins A, B, C, D, and E. See, for example, Alfenore et al., Improving ethanol production and viability of Saccharomyces cerevisiae by a vitamin feeding strategy during fed-batch process, Springer-Verlag (2002), 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., arabinol, n-butanol, isobutanol, ethanol, glycerol, methanol, ethylene glycol, 1,3-propanediol [propylene glycol], butanediol, glycerin, sorbitol, and xylitol); an alkane (e.g., pentane, hexane, heptane, octane, nonane, decane, undecane, and dodecane), a cycloalkane (e.g., cyclopentane, cyclohexane, cycloheptane, and cyclooctane), an alkene (e.g., pentene, hexene, heptene, and octene); an amino acid (e.g., aspartic acid, glutamic acid, glycine, lysine, serine, and threonine); a gas (e.g., methane, hydrogen (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, 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 hydroxylic moieties. In a more preferred aspect, the alcohol is n-butanol. In another more preferred aspect, the alcohol is isobutanol. In another more preferred aspect, the alcohol is ethanol. In another more preferred aspect, the alcohol is methanol. In another more preferred aspect, the alcohol is arabinol. 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 xylitol. In another more preferred aspect, the alcohol is glycerol. In another more preferred aspect, the alcohol is 1,3-propanediol. In another more preferred aspect, the alcohol is sorbitol. In another more preferred aspect, the alcohol is xylitol. See, for example, Gong, C. S., Cao, N. J., Du, J., and Tsao, G. T., 1999, Ethanol production from renewable resources, in Advances in Biochemical Engineering/Biotechnology, Schepers, T., ed., Springer-Verlag Berlin Heidelberg, Germany, 65: 207-241; Silveira, M. M., and Jonas, R., 2002, The biotechnological production of sorbitol, Appl. Microbiol. Biotechnol. 59: 400-408; Nigam, P., and Singh, D., 1995, Processes for fermentative production of xylitol - a sugar substitute, Process Biochemistry 30 (2): 117-124; Ezeji, T. C., Qureshi, N. and Blaschek, H. P., 2003, Production of acetone, butanol and ethanol by Clostridium beijerinckii BA101 and

In 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. Kiriyama, 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 acetonic 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.

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

**Examples**

**Example 1: Pretreatment of corn stover**

Corn stover was pretreated at the U.S. Department of Energy National Renewable Energy Laboratory (NREL) using 1.4% (w/v) sulfuric acid for 8 minutes at 165°C and 107 psi.
The water-insoluble solids in the pretreated corn stover contained 57.5% cellulose, 4.6% hemicellulose, and 28.4% lignin. Cellulose and hemicellulose were determined by a two-stage sulfuric acid hydrolysis with subsequent analysis of sugars by high performance liquid chromatography using NREL Standard Analytical Procedure #002. Lignin was determined gravimetrically after hydrolyzing the cellulose and hemicellulose fractions with sulfuric acid using NREL Standard Analytical Procedure #003.

The pretreated corn stover (PCS) was milled (dry weight 32.35%) in a Cosmos ICMG 40 wet multi utility grinder (EssEmm Corporation, Tamil Nadu, India), and then adjusted to pH 5.0 by repeated addition of 10 N NaOH in aliquots of a few milliliters, followed by thorough mixing and incubation at room temperature for approximately 1 hour. The pH was confirmed after overnight incubation at 4°C, and the pH-adjusted corn stover was autoclaved for 20 minutes at approximately 120°C, and then stored at 4°C to minimize the risk of microbial contamination. The dry weight of the pretreated corn stover was 33% TS (total solids), which was confirmed before each use.

**Example 2: Preparation of phosphoric acid swollen cellulose (PASC)**

A 1% phosphoric acid swollen cellulose (PASC) slurry was prepared from AVICEL® PH101 (Sigma-Aldrich, St. Louis, MO, USA) using the protocol described by Zhang et al., 2006, *Biomacromolecules* 7: 644-648.

**Example 3: Hydrolysis assay**

The effect of a chitin binding protein on the cellulosolytic activity of a cellulase preparation is evaluated according to the procedures described below.

A blend of an *Aspergillus aculeatus* GH10 xylanase (WO 94/021785) and a *Trichoderma reesei* cellulase preparation containing *Aspergillus fumigatus* beta-glucosidase (WO 2005/047499) and *Thermoascus aurantiacus* GH61A polypeptide (WO 2005/074656) available from Novozymes A/S, Bagsvaerd, Denmark, is used as the cellulase preparation. The cellulase preparation is designated herein in the Examples as "*Trichoderma reesei* cellulase composition".

The hydrolysis of PCS is conducted using 2.0 ml deep-well plates (Axygen Scientific, Union City, CA, USA) in a total reaction volume of 1.0 ml. Each hydrolysis is performed with 50 mg of PCS (total insoluble solids; 28.8 mg of cellulose) per ml of 50 mM sodium acetate pH 5.0 buffer containing the *T. reesei* cellulase composition at 2 mg protein per gram of cellulose, plus 1 mM manganese sulfate with and without a chitin binding protein at 0.2 or 1 mg per g cellulose. The chitin binding protein and manganese sulfate are preincubated for 10 minutes at 23°C before mixing with the *T. reesei* cellulase composition, PCS, and buffer. The plate is then sealed using an ALPS-300™ or ALPS-3000™ plate heat sealer (Abgene,
Epsom, United Kingdom), mixed thoroughly, and incubated at 50°C for 1-7 days in an Isotemp Plus incubator (Thermo Fisher Scientific Inc., Waltham, MA, USA). All experiments are performed at least in triplicate.

The hydrolysis of PASC is conducted as described as above, with the exception of using 5 mg of PASC per ml containing no T. reesei cellulase composition, with or without 10 mg of T. aurantiacus GH61A polypeptide and/or 10 mg of B. licheniformis chin binding protein per gram of cellulose, with or without 5 mM pyrogallol, 1 mM manganese sulfate, for 3 days.

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

Percent conversion is calculated based on the mass ratio of solubilized glucosyl units to the initial mass of insoluble cellulose. Only glucose and cellobiose are measured for soluble sugars, as cellodextrins longer than cellobiose are present in negligible concentrations (due to enzymatic hydrolysis). The extent of total cellulose conversion is calculated using the Equation 1:

\[ \text{Percent Conversion} = \frac{\text{Total Sugar Concentration}}{\text{Initial Sugar Concentration}} \times 100 \] (Equation 1)

The 1.111 and 1.053 factors for glucose and cellobiose, respectively, take into account the increase in mass when the glucosyl units in cellulose (average molecular mass of 162 daltons) are converted to glucose (molecular mass of 180 daltons) or cellobiose glucosyl units (average molecular mass of 171 daltons).

Example 4: Preparation of Thermoascus aurantiacus GH61A polypeptide having cellulolytic enhancing activity

Thermoascus aurantiacus GH61A polypeptide was recombinantly produced in Aspergillus oryzae JaL250 according to WO 2005/074656. The recombinantly produced T.
**Example 5: Effect of a chitin binding protein on hydrolysis of PCS by the *Trichoderma reesei* cellulase composition**

The effect of a chitin binding protein on the hydrolysis of PCS by the *T. reesei* cellulase composition is determined using the experimental conditions and procedures described in Example 3.

The effect of the chitin binding protein on hydrolysis of PCS by the *T. reesei* cellulase composition is quantified by determining the ratio of percent conversion of the cellulosic material in the presence of the chitin binding protein to the percent conversion of PCS in the absence of chitin binding protein as shown in Equation 2:

\[
CBP \text{ enhancement effect} = \frac{\% \text{ conversion}^\wedge_{\text{CBP}}}{\% \text{ conversion}^\wedge_{\text{no CBP}}} \quad \text{(Equation 2)}
\]

Stimulation of hydrolysis by the chitin binding protein yields a ratio > 1; inhibition of hydrolysis yields a ratio < 1, and no effect on hydrolysis yields a ratio = 1.

**Example 6: Effect of a chitin binding protein on degradation of PASC**

The effect of a chitin binding protein on degrading PASC is determined using the experimental conditions and procedures described in Example 3.

A CPB-GH61 synergistic effect calculated according to Equation 3.

\[
\text{CBP-GH61 synergistic effect} = \frac{\% \text{ conversion}^\wedge_{\text{CBP}} \cdot \% \text{ conversion}^\wedge_{\text{GH61}}}{\% \text{ conversion}^\wedge_{\text{CBP}} + \% \text{ conversion}^\wedge_{\text{GH61}}} \quad \text{(Equation 3)}
\]

Synergism between a chitin binding protein and a GH61 polypeptide yields a ratio > 1; additiveness yields a ratio = 1, and inhibition yields a ratio < 1.
The present invention is further described by the following numbered paragraphs:


[2] The method of paragraph 1, wherein the chitin binding protein is selected from the group consisting of: (a) a chitin binding protein having at least 60% sequence identity to the full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof; (b) a chitin binding protein encoded by a polynucleotide that hybridizes under at least medium-high stringency conditions with the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; (c) a chitin binding protein encoded by a polynucleotide having at least 60% sequence identity to the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; (d) a variant of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof, comprising a substitution, deletion, and/or insertion at one or more positions; and (e) a fragment of the chitin binding protein of (a), (b), (c), or (d) that has chitin binding activity.

[3] The method of paragraph 2, wherein the chitin binding protein has at least 60%, 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 full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof.

[4] The method of paragraph 2, wherein the chitin binding protein is encoded by a polynucleotide that hybridizes under medium-high stringency conditions, high stringency conditions, or very high stringency conditions with the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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:
NO: 19, SEQ ID NO: 21, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; or the full-length complement thereof.

[5] The method of paragraph 2, wherein the chitin binding protein is encoded by a polynucleotide having at least 60%, 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 full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof.

[6] The method of paragraph 2, wherein the chitin binding protein comprises or consists of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof.

[7] The method of paragraph 2, wherein the chitin binding protein comprises or consists of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof.

[8] The method of paragraph 2, wherein the chitin binding protein is a variant of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof, comprising a substitution, deletion, and/or insertion at one or more positions.

[9] The method of paragraph 2, wherein the chitin binding protein is a fragment of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof; wherein the fragment has chitin binding activity.

[10] The method of any of paragraphs 1-9, wherein the cellulosic material is pretreated.

[11] The method of any of paragraphs 1-10, wherein the cellulosic material is treated with the enzyme composition in the presence of the chitin binding protein and a GH61 polypeptide having cellulytic enhancing activity.

[12] The method of any of paragraphs 1-11, wherein the enzyme composition comprises one or more enzymes selected from the group consisting of a cellulase, a GH61 polypeptide, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.
[13] The method of paragraph 12, wherein the cellulase is one or more enzymes selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase.

[14] The method of paragraph 12, 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.


[16] The method of paragraph 15, wherein the degraded cellulosic material is a sugar.

[17] The method of paragraph 16, wherein the sugar is selected from the group consisting of glucose, xylose, mannose, galactose, and arabinose.

[18] The method of any of paragraphs 1-17, wherein the enzyme composition and/or the chitin binding protein are in the form of a fermentation broth with or without cells.

[19] A method for producing a fermentation product, comprising: (a) saccharifying a cellulosic material with an enzyme composition in the presence of a chitin binding protein; (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.

[20] The method of paragraph 19, wherein the chitin binding protein is selected from the group consisting of: (a) a chitin binding protein having at least 60% sequence identity to the full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof; (b) a chitin binding protein encoded by a polynucleotide that hybridizes under at least medium-high stringency conditions with the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; or the full-length complement thereof; (c) a chitin binding protein encoded by a polynucleotide having at least 60% sequence identity to the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23, or the CBM33 coding sequence thereof; (d) a variant of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24, or the CBM33 thereof, comprising a substitution, deletion,
and/or insertion at one or more positions; and (e) a fragment of the chitin binding protein of
(a), (b), (c), or (d) that has chitin binding activity.

[21] The method of paragraph 20, wherein the chitin binding protein has at least 60%,
5 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 full-length or
mature chitin binding protein of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8,
10 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, or SEQ ID NO: 24; or the CBM33 thereof.

[22] The method of paragraph 20, wherein the chitin binding protein is encoded by a
polynucleotide that hybridizes under medium-high stringency conditions, high stringency
conditions, or very high stringency conditions with the full-length or mature chitin binding
protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7,
15 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; or the
full-length complement thereof.

[23] The method of paragraph 20, wherein the chitin binding protein is encoded by a
polynucleotide having at least 60%, 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 full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1,
20 SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or
the CBM33 coding sequence thereof.

[24] The method of paragraph 20, wherein the chitin binding protein comprises or
consists of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10,
25 SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID
NO: 22, or SEQ ID NO: 24; or the CBM33 thereof.

[25] The method of paragraph 20, wherein the chitin binding protein comprises or
consists of the mature chitin binding protein of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6,
30 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, or SEQ ID NO: 24; or the CBM33 thereof.

[26] The method of paragraph 20, wherein the chitin binding protein is a variant of the
mature chitin binding protein of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8,
35 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, or SEQ ID NO: 24, or the CBM33 thereof, comprising a substitution, deletion, and/or insertion at one or more positions.

[27] The method of paragraph 20, wherein the chitin binding protein is a fragment of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof; wherein the fragment has chitin binding activity.

[28] The method of any of paragraphs 19-27, wherein the cellulosic material is pretreated.

[29] The method of any of paragraphs 19-28, wherein the cellulosic material is treated with the enzyme composition in the presence of the chitin binding protein and a GH61 polypeptide having cellulytic enhancing activity.

[30] The method of any of paragraphs 19-29, wherein the enzyme composition comprises one or more enzymes selected from the group consisting of a cellulase, a GH61 polypeptide, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

[31] The method of paragraph 30, wherein the cellulase is one or more enzymes selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase.

[32] The method of paragraph 30, 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.

[33] The method of any of paragraphs 19-32, wherein steps (a) and (b) are performed simultaneously in a simultaneous saccharification and fermentation.

[34] The method of any of paragraphs 19-33, wherein the fermentation product is an alcohol, an organic acid, a ketone, an amino acid, an alkane, a cycloalkane, an alkene, isoprene, polyketide, or a gas.

[35] The method of any of paragraphs 19-34, wherein the enzyme composition and/or the chitin binding protein are in the form of a fermentation broth with or without cells.

[36] 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 a chitin binding protein.

[37] The method of paragraph 36, wherein the chitin binding protein is selected from the group consisting of: (a) a chitin binding protein having at least 60% sequence identity to the full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof; (b) a
chitin binding protein encoded by a polynucleotide that hybridizes under at least medium-high stringency conditions with the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; or the full-length complement thereof; (c) a chitin binding protein encoded by a polynucleotide having at least 60% sequence identity to the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23, or the CBM33 coding sequence thereof; (d) a variant of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24, or the CBM33 thereof, comprising a substitution, deletion, and/or insertion at one or more positions; and (e) a fragment of the chitin binding protein of (a), (b), (c), or (d) that has chitin binding activity.

[38] The method of paragraph 37, wherein the chitin binding protein has at least 60%, 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 full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof.

[39] The method of paragraph 37, wherein the chitin binding protein is encoded by a polynucleotide that hybridizes under medium-high stringency conditions, high stringency conditions, or very high stringency conditions with the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; or the full-length complement thereof.

[40] The method of paragraph 37, wherein the chitin binding protein is encoded by a polynucleotide having at least 60%, 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 full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or
the CBM33 coding sequence thereof.

[41] The method of paragraph 37, wherein the chitin binding protein comprises or
consists of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof.

[42] The method of paragraph 37, wherein the chitin binding protein comprises or
consists of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof

[43] The method of paragraph 37, wherein the chitin binding protein is a variant of the
mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof, comprising a
substitution, deletion, and/or insertion at one or more positions.

[44] The method of paragraph 37, wherein the chitin binding protein is a fragment of
SEQ ID NO: 2, 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, or
SEQ ID NO: 24; or the CBM33 thereof; wherein the fragment has chitin binding activity.

[45] The method of any of paragraphs 36-44, wherein the fermenting of the cellulosic
material produces a fermentation product.

[46] The method of paragraph 45, further comprising recovering the fermentation
product from the fermentation.

[47] The method of any of paragraphs 36-46, wherein the cellulosic material is
pretreated before saccharification.

[48] The method of any of paragraphs 36-47, wherein the cellulosic material is
treated with the enzyme composition in the presence of the chitin binding protein and a
GH61 polypeptide having cellulosytic enhancing activity.

[49] The method of any of paragraphs 36-48, wherein the enzyme composition
comprises one or more enzymes selected from the group consisting of a cellulase, a GH61
polypeptide, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a
pectinase, a peroxidase, a protease, and a swollenin.

[50] The method of paragraph 49, wherein the cellulase is one or more enzymes
selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-
glucosidase.

[51] The method of paragraph 49, 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.

[52] The method of any of paragraphs 45-51, wherein the fermentation product is an alcohol, an organic acid, a ketone, an amino acid, an alkane, a cycloalkane, an alkene, isoprene, polyketide, or a gas.

[53] The method of any of paragraphs 36-52, wherein the enzyme composition and/or the chitin binding protein are in the form of a fermentation broth with or without cells.

[54] A whole broth formulation, cell culture composition, or enzyme composition comprising a chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

[55] The composition of paragraph 54, which further comprises one or more enzymes selected from the group consisting of a cellulase, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

[56] The composition of paragraph 55, wherein the cellulase is one or more enzymes selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase.

[57] The composition of paragraph 55, 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.

[58] A whole broth formulation, cell culture composition, or enzyme composition comprising a chitin binding protein and one or more enzymes.

[59] The composition of paragraph 58, which the one or more enzymes are selected from the group consisting of a cellulase, a GH61 polypeptide, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

[60] The composition of paragraph 59, wherein the cellulase is one or more enzymes selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase.

[61] The composition of paragraph 59, 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.

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.

Various references are cited herein, the disclosures of which are incorporated by reference in their entireties.
Claims

What is claimed is:

1. A method for degrading or converting a cellulosic material, comprising: treating the cellulosic material with an enzyme composition in the presence of a chitin binding protein.

2. The method of claim 1, wherein the chitin binding protein is selected from the group consisting of:
   (a) a chitin binding protein having at least 60% sequence identity to the full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof;
   (b) a chitin binding protein encoded by a polynucleotide that hybridizes under at least medium-high stringency conditions with the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; or the full-length complement thereof;
   (c) a chitin binding protein encoded by a polynucleotide having at least 60% sequence identity to the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof;
   (d) a variant of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24, or the CBM33 thereof, comprising a substitution, deletion, and/or insertion at one or more positions; and
   (e) a fragment of the chitin binding protein of (a), (b), (c), or (d) that has chitin binding activity.

3. The method of claim 2, wherein the chitin binding protein comprises or consists of SEQ ID NO: 2, 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, or SEQ ID NO: 24; the mature chitin binding protein thereof; or the CBM33 thereof.
4. The method of any of claims 1-3, wherein the cellulosic material is treated with the enzyme composition in the presence of the chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

5. The method of any of claims 1-4, wherein the enzyme composition comprises one or more enzymes selected from the group consisting of a cellulase, a GH61 polypeptide, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

6. The method of any of claims 1-5, further comprising recovering the degraded cellulosic material.

7. The method of claim 6, wherein the degraded cellulosic material is a sugar.

8. A method for producing a fermentation product, comprising: (a) saccharifying a cellulosic material with an enzyme composition in the presence of a chitin binding protein; (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.

9. The method of claim 8, wherein the chitin binding protein is selected from the group consisting of:

   (a) a chitin binding protein having at least 60% sequence identity to the full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof;

   (b) a chitin binding protein encoded by a polynucleotide that hybridizes under at least medium-high stringency conditions with the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; or the full-length complement thereof;

   (c) a chitin binding protein encoded by a polynucleotide having at least 60% sequence identity to the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof;
(d) a variant of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24, or the CBM33 thereof, comprising a substitution, deletion, and/or insertion at one or more positions; and

(e) a fragment of the chitin binding protein of (a), (b), (c), or (d) that has chitin binding activity.

10. The method of claim 9, wherein the chitin binding protein comprises or consists of SEQ ID NO: 2, 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, or SEQ ID NO: 24; the mature chitin binding protein thereof; or the CBM33 thereof.

11. The method of any of claims 8-10, wherein the cellulosic material is treated with the enzyme composition in the presence of the chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

12. The method of any of claims 8-11, wherein the enzyme composition comprises one or more enzymes selected from the group consisting of a cellulase, a GH61 polypeptide, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

13. 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 a chitin binding protein.

14. The method of claim 13, wherein the chitin binding protein is selected from the group consisting of:

(a) a chitin binding protein having at least 60% sequence identity to the full-length or mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24; or the CBM33 thereof;

(b) a chitin binding protein encoded by a polynucleotide that hybridizes under at least medium-high stringency conditions with the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23; or the CBM33 coding sequence thereof; or the full-length complement thereof;
(c) a chitin binding protein encoded by a polynucleotide having at least 60% sequence identity to the full-length or mature chitin binding protein coding sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ 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, or SEQ ID NO: 23, or the CBM33 coding sequence thereof;

(d) a variant of the mature chitin binding protein of SEQ ID NO: 2, 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, or SEQ ID NO: 24, or the CBM33 thereof, comprising a substitution, deletion, and/or insertion at one or more positions; and

(e) a fragment of the chitin binding protein of (a), (b), (c), or (d) that has chitin binding activity.

15. The method of claim 14, wherein the chitin binding protein comprises or consists of SEQ ID NO: 2, 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, or SEQ ID NO: 24; the mature chitin binding protein thereof; or the CBM33 thereof.

16. The method of any of claims 13-15, wherein the fermenting of the cellulosic material produces a fermentation product.

17. The method of claim 16, further comprising recovering the fermentation product from the fermentation.

18. The method of any of claims 13-17, wherein the cellulosic material is treated with the enzyme composition in the presence of the chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

19. The method of any of claims 13-18, wherein the enzyme composition comprises one or more enzymes selected from the group consisting of a cellulase, a GH61 polypeptide, a hemicellulase, an esterase, an expansin, a laccase, a ligninolytic enzyme, a pectinase, a peroxidase, a protease, and a swollenin.

20. A whole broth formulation, cell culture composition, or enzyme composition comprising a chitin binding protein and a GH61 polypeptide having cellulolytic enhancing activity.

21. A whole broth formulation, cell culture composition, or enzyme composition comprising a chitin binding protein and one or more enzymes.
# A. CLASSIFICATION OF SUBJECT MATTER

INV. C12P7/10 C12P19/02 C12P19/04 C12P19/14

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)
C12P

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, EMBASE, WPI Data

# C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td></td>
<td>1-21</td>
</tr>
</tbody>
</table>

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 application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) on 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

Date of the actual completion of the international search: 20 July 2012

Date of mailing of the international search report: 16/10/2012

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: Mateo Rosel 1, A

Form PCT/ISA210 (second sheet) (April 2005)
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>WO 2009/018537 A2 (DYADIC INTERNATIONAL INC [US]; NEDWIN GLENN [US]; GUSAKOV ALEXANDER VA) 5 February 2009 (2009-02-05)</td>
<td>1</td>
</tr>
<tr>
<td>Y</td>
<td>XW 2009/018537 A2 (DYADIC INTERNATIONAL INC [US]; NEDWIN GLENN [US]; GUSAKOV ALEXANDER VA) 5 February 2009 (2009-02-05)</td>
<td>1-21</td>
</tr>
<tr>
<td>Y</td>
<td>CHU H H ET AL: &quot;A Bacterius amyloliquefaciens ChbB protein binds [beta]- and [alpha]-chitin and has homologues in related strains&quot;, MICROBIOLOGY 2001 GB, vol. 147, no. 7, 2001, pages 1793-1803, XP002680368, ISSN: 1350-0872 abstract page 1799, left-hand column, last paragraph - page 1800, right-hand column, paragraph 1; figures 3, 5, 8 page 17</td>
<td>1-21</td>
</tr>
</tbody>
</table>
INTERNATIONAL SEARCH REPORT

Box No. II  Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. □ Claims Nos.:
   because they relate to subject matter not required to be searched by this Authority, namely:

2. □ Claims Nos.:
   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. □ Claims Nos.:
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III  Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. □ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. □ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. □ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. □ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

   I-21 (partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet (2)) (April 2005)
This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: l-21 (partially)

A method for degrading or converting a cellulose material comprising: treating the cellulose material with an enzyme composition in the presence of: (a) a chitin binding protein having at least 60% sequence identity to the full length or mature chitin binding protein of SEQ. ID. N.2; (b) chitin binding protein ns encoded by polynucleotide sequences which hybridize to SEQ. ID. N.1; (c) chitin binding protein ns encoded by polynucleotide sequences having at least 60% identity to the full length or mature chitin binding protein of SEQ. ID. N.1; (d) a variant of the mature chitin binding protein of SEQ. ID. N.2; (e) a fragment of the chitin binding protein of a), b), c) or d) that has chitin binding activity.

2-13. Claims: l-21 (partially)

A method for degrading or converting a cellulose material comprising: treating the cellulose material with an enzyme composition in the presence of: (a) a chitin binding protein having at least 60% sequence identity to the full length or mature chitin binding protein of SEQ. ID. N.4, SEQ. ID. N.6, SEQ. ID. N.8, SEQ. ID. N.10, SEQ. ID. N.12, SEQ. ID. N.14, SEQ. ID. N.16, SEQ. ID. N.18, SEQ. ID. N.20, SEQ. ID. N.22, SEQ. ID. N.24 or the CBM; (b) chitin binding protein ns encoded by polynucleotide sequences which hybridize to SEQ. ID. N.3, SEQ. ID. N.5, SEQ. ID. N.7, SEQ. ID. N.9, SEQ. ID. N.11, SEQ. ID. N.13, SEQ. ID. N.15, SEQ. ID. N.17, SEQ. ID. N.19, SEQ. ID. N.21, SEQ. ID. N.23 or the CBM thereof; (c) chitin binding protein ns encoded by polynucleotide sequences having at least 60% identity to the full length or mature chitin binding protein of SEQ. ID. N.1; (d) a variant of the mature chitin binding protein of SEQ. ID. N.4, SEQ. ID. N.6, SEQ. ID. N.8, SEQ. ID. N.10, SEQ. ID. N.12, SEQ. ID. N.14, SEQ. ID. N.16, SEQ. ID. N.18, SEQ. ID. N.20, SEQ. ID. N.22, SEQ. ID. N.24 or the CBM; (e) a fragment of the chitin binding protein of a), b), c) or d) that has chitin binding activity.
<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family member(s)</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td>US 2007218046 A1</td>
<td>20-09-2007</td>
<td>NONE</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2012030838 A1</td>
<td>02-02-2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2012036599 A1</td>
<td>09-02-2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2009018537 A2</td>
<td>05-02-2009</td>
</tr>
</tbody>
</table>