



(43) International Publication Date
11 December 2014 (11.12.2014)

(10) International Publication Number
WO 2014/197705 A1

- (51) **International Patent Classification:**
C12N 9/22 (2006.01) *C12P 7/06* (2006.01)
- (21) **International Application Number:**
PCT/US2014/041113
- (22) **International Filing Date:**
5 June 2014 (05.06.2014)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/832,470 7 June 2013 (07.06.2013) US
- (71) **Applicant:** NOVOZYMES A/S [DK/DK]; Krogshoejvej 36, DK-2880 Bagsvaerd (DK).
- (72) **Inventor; and**
- (71) **Applicant** (for BW only): SOONG, Chee-Leong [MY/US]; 12520 Honeychurch St., Raleigh, North Carolina 27614 (US).
- (72) **Inventors:** HARRIS, Paul; 11927 342nd Avenue NE, Carnation, Washington 98014 (US). KIEMER, Lars; Oestervej 43, DK-2750 Ballerup (DK).
- (74) **Agents:** KRENICKY, Michael W. et al.; Novozymes North America, Inc., 60 E. 42nd St., Suite 700, New York, New York 10165 (US).
- (81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,

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

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

Declarations under Rule 4.17:

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

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) **Title:** METHOD AND COMPOSITIONS FOR ENHANCING AN X143 POLYPEPTIDE HAVING RETROGRADED STARCH DEGRADING ACTIVITY AND USES THEREOF

(57) **Abstract:** The present invention relates to a composition comprising: (a) an X143 polypeptide having retrograded starch degrading activity and (b) a sulfur-containing compound, and/or a heterocyclic compound (reducing agent), wherein the combination of the X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, enhances degradation of retrograded starch material by the X143 polypeptide. The present invention further relates to a method for degrading or converting a retrograded starch material, comprising: treating the retrograded starch material with an enzyme composition in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound and/or the heterocyclic compound (reducing agent), and/or the cellobiose dehydrogenase enhances degradation of the retrograded starch material by the enzyme composition.



WO 2014/197705 A1

METHOD AND COMPOSITIONS FOR ENHANCING AN X143 POLYPEPTIDE HAVING RETROGRADED STARCH DEGRADING ACTIVITY AND USES THEREOF

Reference to a Sequence Listing

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

Background of the Invention

Field of the Invention

10 The present invention relates to a method of enhancing an X143 polypeptide having retrograded starch degrading activity by the addition of a sulfur-containing compound, a heterocyclic (reducing agent) compound, or a cellobiose dehydrogenase.

 The present invention further relates to compositions comprising an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound, or a
15 heterocyclic (reducing agent) compound, and to methods of using the compositions.

Description of the Related Art

 Retrograded starch or resistant starch (RS) potentially forms during the heating-cooling cycles in typical conventional corn to ethanol processes or food industry processes. According
20 to literature these starch structures are thermally very stable and does not break down in normal cooking operations and temperatures (Haralampu, 2000, Resistant starch—a review of the physical properties and biological impact of RS3. *Carbohydrate Polymers* 41: 285–292; Szymonska et al., 2003, Modification of granular potato starch by multiple deep-freezing and thawing, *Carbohydrate Polymers* 52: 1-10; Karim, 2000, Methods for the study of starch
25 retrogradation, *Food Chemistry* 71: 9-36). In many cases, retrograded starch or resistant starch is not easily accessible to hydrolytic enzymes. The X143 enzyme can cleave retrograded starch (formed during liquefaction process) and will release oligosaccharides and sugars and ultimately increase ethanol yield in conventional 1st gen biofuel. A X143 polypeptide was described in WO 2010/059413 as SEQ ID NO: 4. The X143 enzyme can potentially be used in
30 the food (baking) or brewery industry to remove retrograded starch to improve foods property and smooth processing.

 It would be advantageous in the art to improve the ability of X143 polypeptides to degrade retrograded starch.

 The present invention relates to compositions comprising an X143 polypeptide and a
35 sulfur-containing compound or a reducing agent compound, and to methods of using the compositions.

Summary of the Invention

In a first aspect the present invention relates to a composition comprising: (a) an X143 polypeptide having retrograded starch degrading activity and (b) a sulfur-containing compound, and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the X143 polypeptide having retrograded starch degrading activity and the sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, enhances degradation of retrograded starch material by the X143 polypeptide.

In a second aspect the invention relates to a method for degrading a retrograded starch material, comprising: treating the retrograded starch material with an enzyme composition in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition.

In a third aspect the present invention relates to a method of saccharifying a starch containing material, comprising:

(a) liquefying the starch containing material in the presence of at least an alpha amylase;

(b) saccharifying the starch containing material in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound, and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, and at least a glucoamylase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or the

cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition.

In a fourth aspect the present invention relates to a method for producing a fermentation product, comprising:

(a) liquefying a starch containing material in the presence of at least an alpha amylase,

(b) saccharifying the starch containing material in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, and at least a glucoamylase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or the cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition;

(c) fermenting the saccharified starch material with one or more fermenting microorganisms to produce the fermentation product.

Definitions

Alkyl: The term "alkyl," by itself or as part of another substituent, means, unless otherwise stated, a fully saturated straight-chain (linear; unbranched) or branched chain, or combination thereof, having the number of carbon atoms specified, if designated (*i.e.* C₁-C₁₀ means one to ten carbons). Examples include, but are not limited to, groups such as methyl, ethyl, n-propyl, isopropyl, n-butyl, t-butyl, isobutyl, sec-butyl, homologs and isomers of, for example, n-pentyl, n-hexyl, n-heptyl, n-octyl, and the like. If no size is designated, the alkyl groups mentioned herein contain 1-20 carbon atoms, typically 1-10 carbon atoms, or 1-8 carbon atoms, or 1-6 carbon atoms, or 1-4 carbon atoms. The term "alkylene" is by itself or in combination with other terms, represents a divalent radical derived from an alkyl, as exemplified, but not limited, by -CH₂CH₂CH₂CH₂-.

Alkenyl: The term "alkenyl" refers to unsaturated aliphatic groups including straight-chain (linear; unbranched), branched-chain groups, and combinations thereof, having the number of carbon atoms specified, if designated, which contain at least one double bond (-C=C-). All double bonds may be independently either (E) or (Z) geometry, as well as mixtures thereof. Examples of alkenyl groups include, but are not limited to, -CH₂-CH=CH-CH₃; -CH=CH-CH=CH₂ and -CH₂-CH=CH-CH(CH₃)-CH₂-CH₃. If no size is

designated, the alkenyl groups mentioned herein contain 2-20 carbon atoms, typically 2-10 carbon atoms, or 2-8 carbon atoms, or 2-6 carbon atoms, or 2-4 carbon atoms. The term "alkenylene" is by itself or in combination with other terms, represents a divalent radical derived from an alkenyl, as exemplified, but not limited, by $-\text{CH}_2\text{CHCHCH}_2-$.

5 **Alkynyl:** The term "alkynyl" refers to unsaturated aliphatic groups including straight-chain (linear; unbranched), branched-chain groups, and combinations thereof, having the number of carbon atoms specified, if designated, which contain at least one carbon-carbon triple bond ($-\text{C}\equiv\text{C}-$). Examples of alkynyl groups include, but are not limited to, $-\text{CH}_2-\text{C}\equiv\text{C}-\text{CH}_3$; $-\text{C}\equiv\text{C}-\text{C}\equiv\text{CH}$ and $-\text{CH}_2-\text{C}\equiv\text{C}-\text{CH}(\text{CH}_3)-\text{CH}_2-\text{CH}_3$. If no size is designated, the
10 alkynyl groups mentioned herein contain 2-20 carbon atoms, typically 2-10 carbon atoms, or 2-8 carbon atoms, or 2-6 carbon atoms, or 2-4 carbon atoms. The term "alkynylene" is by itself or in combination with other terms, represents a divalent radical derived from an alkynyl, as exemplified, but not limited, by $-\text{CH}_2\text{CCCH}_2-$.

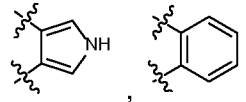
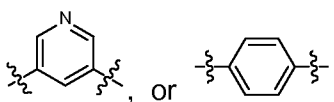
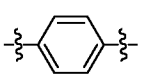
15 **Allelic variant:** The term "allelic variant" means any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in polymorphism within populations. Gene mutations can be silent (no change in the encoded 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.

20 **Amino acid moiety:** With respect to the sulfur-containing compounds described herein, the term "amino acid moiety" refers to a radical of any of the naturally occurring amino acids, as well as synthetic analogs (e.g., D-stereoisomers of the naturally occurring amino acids, such as D-valine or D-alanine) and derivatives thereof. Amino acids comprise a carbon atom to which is bonded an amino group, a carboxyl group, a hydrogen atom, and a distinctive group referred to
25 as a "side chain". The side chains of naturally occurring amino acids are well known in the art and include, for example, hydrogen (e.g., as in glycine), alkyl (e.g., as in alanine, valine, leucine, isoleucine, proline), substituted alkyl (e.g., as in threonine, serine, methionine, cysteine, aspartic acid, asparagine, glutamic acid, glutamine, arginine, and lysine), aryl alkyl (e.g., as in phenylalanine and tryptophan), substituted aryl alkyl (e.g., as in tyrosine), and
30 heteroarylalkyl (e.g., as in histidine). Unnatural amino acids are also known in the art, as set forth in, for example, Williams (ed.), *Synthesis of Optically Active α -Amino Acids*, Pergamon Press (1989); Evans et al., 1990, *J. Amer. Chern. Soc.* 112: 4011-4030; Pu et al., 1991, *J. Amer. Chern. Soc.* 56: 1280-1283; Williams et al., 1991, *J. Amer. Chern. Soc.* 113: 9276-9286 (and all references cited therein), the contents of which are hereby incorporated herein by
35 reference in its entirety, particularly with respect to the amino acids described therein. The amino acids may include the side chains of known unnatural amino acids as well natural amino acids, unless otherwise stated.

Aralkyl: The term "aralkyl" designates an alkyl-substituted aryl group, where the alkyl portion is attached to the parent structure. Examples are benzyl, phenethyl, and the like. "Heteroaralkyl" designates a heteroaryl moiety attached to the parent structure via an alkyl residue. Examples include furanylmethyl, pyridinylmethyl, pyrimidinylethyl, and the like. Aralkyl and heteroaralkyl also include substituents in which at least one carbon atom of the alkyl group is present in the alkyl group and wherein another carbon of the alkyl group has been replaced by, for example, an oxygen, nitrogen or sulfur atom (e.g., phoxymethyl, 2-pyridylmethoxy, 3-(1-naphthylthio)propyl, and the like).

Aryl: The term "aryl" means, unless otherwise stated, a polyunsaturated, aromatic, hydrocarbon substituent. Aryl may contain additional fused rings (e.g., from 1 to 3 rings), including additionally fused aryl, heteroaryl, cycloalkyl, and/or heterocycloalkyl rings. Examples of aryl groups include, but are not limited to, phenyl, 1-naphthyl, 2-naphthyl, and 4-biphenyl.

Arylene/heteroarylene: The term "arylene" and "heteroarylene" means a divalent radical derived from an aryl and heteroaryl, respectively. Each of the two valencies of arylene

and heteroarylene may be located at any suitable portion of the ring (e.g., , , or ) and may be fused to another ring, as appropriate. Non-limiting

examples of arylene include phenylene, biphenylene, naphthylene, and the like. Examples of heteroarylene groups include, but are not limited to, pyridinylene, oxazolylene, thiazolylene, pyrazolylene, pyranylene, and furanylene.

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

Cellobiose dehydrogenase: The term "cellobiose dehydrogenase" is defined herein as a cellobiose:acceptor 1-oxidoreductase (E.C. 1.1.99.18) that catalyzes the conversion of cellobiose in the presence of an acceptor to cellobiono-1,5-lactone and a reduced acceptor. 2,6-Dichloroindophenol can act as acceptor, as can iron, especially $\text{Fe}(\text{SCN})_3$, molecular oxygen, ubiquinone, or cytochrome C, and likely many other polyphenolics. Substrates of the enzyme include cellobiose, cello-oligosaccharides, lactose, and D-glucosyl-1,4- β -D-mannose, glucose, maltose, mannobiose, thiocellobiose, galactosyl-mannose, xylobiose, xylose. Electron donors are preferably beta-1-4 dihexoses with glucose or mannose at the reducing end, though alpha-1-4 hexosides, hexoses, pentoses, and beta-1-4 pentomers have also been shown to act as substrates for these enzymes (Henriksson *et al*, 1998, *Biochimica et Biophysica Acta* -

Protein Structure and Molecular Enzymology; 1383: 48-54; and Schou *et al*, 1998, *Biochem. J.* 330: 565-571).

Cellobiose dehydrogenases comprise two families, 1 and 2, differentiated by the presence of a cellulose binding motif (CBM). The 3-dimensional structure of cellobiose dehydrogenase features two globular domains, each containing one of two cofactors: a heme or a flavin. The active site lies at a cleft between the two domains. The catalytic cycle of cellobiose dehydrogenase follows an ordered sequential mechanism. Oxidation of cellobiose occurs via 2-electron transfer from cellobiose to the flavin, generating cellobiono-1,5-lactone and reduced flavin. The active FAD is regenerated by electron transfer to the heme group, leaving a reduced heme. The native state heme is regenerated by reaction with the oxidizing substrate at the second active site.

The oxidizing substrate is preferentially iron ferricyanide, cytochrome C, or an oxidized phenolic compound such as dichloroindophenol (DCIP), a substrate commonly used for colorimetric assays. Metal ions and O₂ are also substrates, but for most cellobiose dehydrogenases the reaction rate for these substrates is several orders of magnitude lower than that observed for iron or organic oxidants. Following cellobionolactone release, the product may undergo spontaneous ring-opening to generate cellobionic acid (Hallberg *et al.*, 2003, *J. Biol. Chem.* 278: 7160-7166).

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 usually begins with the ATG start codon or alternative start codons such as GTG and TTG and ends with a stop codon such as TAA, TAG, and TGA. The coding sequence may be a DNA, cDNA, synthetic, or recombinant polynucleotide. In one embodiment the coding sequence consist of positions 1 to 102 joined to positions 162 to 1217 of SEQ ID NO: 1.

Control sequences: The term "control sequences" means all components necessary for the expression of a polynucleotide encoding a polypeptide. Each control sequence may be native or foreign 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.

Cycloalkyl: The term "cycloalkyl" by itself or in combination with other terms, represents, unless otherwise stated, a saturated or unsaturated cyclic non-aromatic hydrocarbon radical (*e.g.*, cyclic versions of alkyl, alkenyl, or alkynyl, or mixtures thereof).

Cycloalkyl may contain additional fused rings (e.g., from 1 to 3 rings), including additionally fused cycloalkyl and/or heterocycloalkyl rings, but excludes additionally fused aryl and/or heteroaryl groups. Examples of cycloalkyl include, but are not limited to, cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, 1-cyclohexenyl, 3-cyclohexenyl, cycloheptyl, norbornyl, and the like. If no size is designated, the alkynyl groups mentioned herein contain 3-9 carbon atoms, typically 3-7 carbon atoms. The term "cycloalkylene" by itself or as part of another substituent means a divalent radical derived from a cycloalkyl, as exemplified, but not limited, by -cyclohexyl-.

Cycloalkyl-alkyl/heterocycloalkyl-alkyl: The terms "cycloalkyl-alkyl" and "heterocycloalkyl-alkyl" designate an alkylsubstituted cycloalkyl group and alkyl-substituted heterocycloalkyl, respectively, where the alkyl moiety is attached to the parent structure. Non-limiting examples include cyclopropylethyl, cyclobutyl-propyl, cyclopentyl-hexyl, cyclohexyl-isopropyl, 1-cyclohexenyl-propyl, 3-cyclohexenyl-t-butyl, cycloheptyl-heptyl, norbornyl-methyl, 1-piperidinyl-ethyl, 4-morpholinyl-propyl, 3-morpholinyl-t-butyl, tetrahydrofuran-2-yl-hexyl, tetrahydrofuran-3-ylisopropyl, and the like. Cycloalkyl-alkyl and heterocycloalkyl-alkyl also include substituents in which at least one carbon atom is present in the alkyl group and wherein another carbon atom of the alkyl group has been replaced by, for example, an oxygen, nitrogen or sulfur atom (e.g., cyclopropoxymethyl, 2-piperidinyloxy-t-butyl, and the like).

Expression: The term "expression" includes any step involved in the production of the 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 additional nucleotides that provide for its expression.

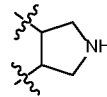
Halogen: The terms "halo" or "halogen," by themselves or as part of another substituent, mean, unless otherwise stated, a fluorine, chlorine, bromine, or iodine atom.

Heteroaryl: The term "heteroaryl" refers to aryl groups (or rings) that contain from one to four annular heteroatoms selected from N, O, and S, wherein the nitrogen and sulfur atoms are optionally oxidized, and the nitrogen atom(s) are optionally quaternized. A heteroaryl group can be attached to the remainder of the molecule at an annular carbon or annular heteroatom.

Heteroaryl may contain additional fused rings (e.g., from 1 to 3 rings), including additionally fused aryl, heteroaryl, cycloalkyl, and/or heterocycloalkyl rings. Non-limiting examples of heteroaryl groups are 1-pyrrolyl, 2-pyrrolyl, 3-pyrrolyl, 3-pyrazolyl, 2-imidazolyl, 4-imidazolyl, pyrazinyl, 2-oxazolyl, 4-oxazolyl, 2-phenyl-4-oxazolyl, 5-oxazolyl, 3-isoxazolyl, 4-isoxazolyl, 5-isoxazolyl, 2-thiazolyl, 4-thiazolyl, 5-thiazolyl, 2-furyl, 3-furyl, 2-thienyl, 3-thienyl, 2-pyridyl, 3-pyridyl, 4-pyridyl, 2-pyrimidyl, 4-pyrimidyl, 5-benzothiazolyl, purinyl, 2-benzimidazolyl, 5-indolyl, 1-isoquinolyl, 5-isoquinolyl, 2-quinoxaliny, 5-quinoxaliny, 3-quinolyl, and 6-quinolyl.

Heterocycloalkyl: The term "heterocycloalkyl," by itself or in combination with other terms, represents a saturated or unsaturated cyclic non-aromatic hydrocarbon radical containing of at least one carbon atom and at least one annular heteroatom selected from the group consisting of O, N, P, Si and S, and wherein the nitrogen and sulfur atoms may optionally be oxidized and the nitrogen heteroatom may optionally be quaternized. The heteroatom(s) O, N, P, S and Si may be placed at any interior position of the heterocycloalkyl group or at the position at which the heterocycloalkyl group is attached to the remainder of the molecule. Heterocycloalkyl may contain additional fused rings (e.g., from 1 to 3 rings), including additionally fused cycloalkyl and/or heterocycloalkyl rings, but excludes additionally fused aryl and/or heteroaryl groups. Examples of heterocycloalkyl include, but are not limited to, thiazolidinonyl, 1-(1,2,5,6-tetrahydropyridyl), 1-piperidinyl, 2-piperidinyl, 3-piperidinyl, 4-morpholinyl, 3-morpholinyl, tetrahydrofuran-2-yl, tetrahydrofuran-3-yl, tetrahydrothien-2-yl, tetrahydrothien-3-yl, 1-piperazinyl, 2-piperazinyl, and the like. The term "heterocycloalkylene" by itself or as part of another substituent means a divalent radical derived from a

heterocycloalkyl, as exemplified, but not limited, by



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

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

Mature polypeptide: The term "mature polypeptide" is defined herein as a polypeptide in its final form following translation and any post-translational modifications, such as N-terminal processing, C-terminal truncation, glycosylation, phosphorylation, etc. In one aspect, the mature polypeptide is amino acids 19 to 385 of SEQ ID NO: 2 based on the SignalP program (Nielsen *et al.*, 1997, *supra*) that predicts amino acids 1 to 18 of SEQ ID NO: 2 is a signal peptide.

Mature polypeptide coding sequence: The term “mature polypeptide coding sequence” is defined herein as a nucleotide sequence that encodes a mature polypeptide having retrograded starch activity. In one aspect, the mature polypeptide coding sequence is nucleotides 55 to 1214 of SEQ ID NO: 1 based on the SignalP program (Nielsen *et al.*, 1997, *supra*) that predicts nucleotides 1 to 54 of SEQ ID NO: 1 encode a signal peptide.

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. The term nucleic acid construct is synonymous with the term “expression cassette” when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present invention.

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

Polypeptide fragment: The term “fragment” means a polypeptide having one or more (e.g., several) amino acids deleted from the amino and/or carboxyl terminus of a mature polypeptide; wherein the fragment has biological activity.

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.

Retrogradation is a reaction that takes place in gelatinized starch when the amylose and amylopectin chains realign themselves, causing the liquid to gel.

When native starch is heated and dissolves in water, the crystalline structure of amylose and amylopectin molecules are lost and they hydrate to form a viscous solution. If the viscous solution is cooled or left at lower temperature for long enough period, the linear molecules, amylose, and linear parts of amylopectin molecules retrograde and rearrange themselves again to a more crystalline structure.

Retrograded starch or resistant Starch (RS) is a less digestible form of starch and is formed by heating, e.g. during liquefaction, and subsequent cooling.

Retrograded starch degrading activity means that the X143 polypeptide is capable of breaking glucosidic bonds in the retrograded starch material. The retrograded starch degrading activity can be assayed as described in the examples.

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

For purposes of the present invention, the degree of sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *J. Mol. Biol.* 48: 443-453) as implemented in the Needle program of the

EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *Trends Genet.* 16: 276-277), preferably version 5.0.0, or later. The optional parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle

5 labeled "longest identity" (obtained using the *-nobrief* option) is used as the percent identity and is calculated as follows:

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

For purposes of the present invention, the degree of sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm

10 (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 optional parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the *-nobrief*

15 option) is used as the percent identity and is calculated as follows:

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

Subsequence: The term "subsequence" means a polynucleotide having one or more (e.g., several) nucleotides deleted from the 5' and/or 3' end of a mature polypeptide coding

20 sequence; wherein the subsequence encodes a fragment having biological activity.

Substituted: The term "substituted" refers to the replacement of one or more (e.g., several) hydrogen atoms of a moiety with a monovalent or divalent radical. "Optionally substituted" indicates that the moiety may be substituted or unsubstituted. A moiety lacking the terms "optionally substituted" and "substituted" is intended an unsubstituted moiety (e.g.,

25 "phenyl" is intended an unsubstituted phenyl unless indicated as a substituted phenyl or an optionally substituted phenyl). Suitable substituent groups for indicated optionally substituted moieties include, for example, hydroxyl, nitro, amino (e.g., -NH_2 or dialkyl amino), imino, cyano, halo (such as F, Cl, Br, I), halo alkyl (such as -CCl_3 or -CF_3), thio, sulfonyl, thioamido, amidino, imidino, oxo, oxamidino, methoxamidino, imidino, guanidino, sulfonamido, carboxyl, formyl,

30 alkyl, alkoxy, alkoxy-alkyl, alkylcarbonyl, alkylcarbonyloxy (-OCOR), aminocarbonyl, arylcarbonyl, aralkylcarbonyl, carbonylamino, heteroarylcarbonyl, heteroaralkyl-carbonyl, alkylthio, amino alkyl, cyanoalkyl, carbamoyl (-NHCOOR- or -OCONHR-), urea (-NHCONHR-), aryl and the like, where R is any suitable group, e.g., alkyl or alkylene. In some embodiments, the optionally substituted moiety is optionally substituted only with select radicals, as described.

35 In some embodiments, the above groups (e.g., alkyl groups) are optionally substituted with, for example, alkyl (e.g., methyl or ethyl), halo alkyl (e.g., -CCl_3 , $\text{-CH}_2\text{CHCl}_3$ or -CF_3), cycloalkyl (e.g., $\text{-C}_3\text{H}_5$, $\text{-C}_4\text{H}_7$, $\text{-C}_5\text{H}_9$), amino (e.g., -NH_2 or dialkyl amino), alkoxy (e.g., methoxy),

heterocycloalkyl (e.g., as morpholine, piperazine, piperidine, azetidine), hydroxyl, and/or heteroaryl (e.g., oxazolyl). Other suitable substituent groups for indicated optionally substituted moieties are described herein. In some embodiments, a substituent group is itself optionally substituted. In some embodiments, a substituent group is not itself substituted. The group substituted onto the substitution group can be, for example, carboxyl, halo, nitro, amino, cyano, hydroxyl, alkyl, alkenyl, alkynyl, alkoxy, aminocarbonyl, -SR, thioamido, -SO₃H, -SO₂R or cycloalkyl, where R is any suitable group, e.g., a hydrogen or alkyl.

When the substituted substituent includes a straight chain group, the substituent can occur either within the chain (e.g., 2-hydroxypropyl, 2-aminobutyl, and the like) or at the chain terminus (e.g., 2-hydroxyethyl, 3-cyanopropyl, and the like). Substituted substituents can be straight chain, branched or cyclic arrangements of covalently bonded carbon or heteroatoms (N, O or S).

Variant: The term “variant” means a polypeptide having cellulolytic enhancing activity comprising an alteration, *i.e.*, a substitution, insertion, and/or deletion of one or more (e.g., several) amino acid residues at one or more (e.g., several) positions. A substitution means a replacement of an amino acid occupying a position with a different amino acid; a deletion means removal of an amino acid occupying a position; and an insertion means adding one or more (e.g., several) amino acids, e.g., 1-5 amino acids, adjacent to an amino acid occupying a position.

Cellulolytic enzyme, cellulolytic composition, or cellulase: The term “cellulolytic enzyme”, “cellulolytic composition”, or “cellulase” means one or more (e.g., several) enzymes that hydrolyze a cellulosic material. Such enzymes include endoglucanase(s), cellobiohydrolase(s), beta-glucosidase(s), or combinations thereof. The two basic approaches for measuring cellulolytic activity include: (1) measuring the total cellulolytic activity, and (2) measuring the individual cellulolytic activities (endoglucanases, cellobiohydrolases, and beta-glucosidases) as reviewed in Zhang *et al.*, Outlook for cellulase improvement: Screening and selection strategies, 2006, *Biotechnology Advances* 24: 452-481. Total cellulolytic activity is usually measured using insoluble substrates, including Whatman №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 №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).

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 Pretreated Corn Stover (“PCS”) (or other pretreated cellulosic material) for 3-7 days at a suitable temperature, e.g., 50°C, 55°C, or 60°C, compared

to a control hydrolysis without addition of cellulolytic enzyme protein. Typical conditions are 1 ml reactions, washed or unwashed PCS, 5% insoluble solids, 50 mM sodium acetate pH 5, 1 mM MnSO₄, 50°C, 55°C, or 60°C, 72 hours, sugar analysis by AMINEX® HPX-87H column (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

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, 1991, A classification of glycosyl hydrolases based on amino-acid sequence similarities, *Biochem. J.* 280: 309-316, and Henrissat and Bairoch, 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.

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

The GH61 polypeptide 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.

Beta-glucosidase: The term “beta-glucosidase” means a beta-D-glucoside glucohydrolase (E.C. 3.2.1.21) that catalyzes the hydrolysis of terminal non-reducing beta-D-glucose residues with the release of beta-D-glucose.

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

Cellobiohydrolase: The term “cellobiohydrolase” means a 1,4-beta-D-glucan cellobiohydrolase (E.C. 3.2.1.91) that catalyzes the hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, cellobiosaccharides, or any beta-1,4-linked glucose containing polymer, releasing cellobiose from the reducing or non-reducing ends of the chain (Teeri, 1997, Crystalline cellulose degradation: New insight into the function of cellobiohydrolases, *Trends in Biotechnology* 15: 160-167; Teeri *et al.*, 1998, *Trichoderma reesei* cellobiohydrolases: why so efficient on crystalline cellulose?, *Biochem. Soc. Trans.* 26: 173-178).

Cellobiohydrolase activity is determined according to the procedures described by Lever *et al.*, 1972, *Anal. Biochem.* 47: 273-279; van Tilbeurgh *et al.*, 1982, *FEBS Letters*, 149: 152-156; van Tilbeurgh and Claeysens, 1985, *FEBS Letters*, 187: 283-288; and Tomme *et al.*, 1988, *Eur. J. Biochem.* 170: 575-581. In the present invention, the Tomme *et al.* method can be used to determine cellobiohydrolase activity.

Endoglucanase: The term “endoglucanase” means an endo-1,4-(1,3;1,4)-beta-D-glucan 4-glucanohydrolase (E.C. 3.2.1.4) that catalyzes endohydrolysis of 1,4-beta-D-glycosidic linkages in cellulose, cellulose derivatives (such as carboxymethyl cellulose and hydroxyethyl cellulose), lichenin, beta-1,4 bonds in mixed beta-1,3 glucans such as cereal beta-D-glucans or xyloglucans, and other plant material containing cellulosic components.

Endoglucanase activity can be determined by measuring reduction in substrate viscosity or increase in reducing ends determined by a reducing sugar assay (Zhang *et al.*, 2006, *Biotechnology Advances* 24: 452-481). For purposes of the present invention, endoglucanase activity is determined using carboxymethyl cellulose (CMC) as substrate according to the procedure of Ghose, 1987, *Pure and Appl. Chem.* 59: 257-268, at pH 5, 40°C.

Hemicellulolytic enzyme or hemicellulase: The term “hemicellulolytic enzyme” or “hemicellulase” means one or more (e.g., several) enzymes that hydrolyze a hemicellulosic material. See, for example, Shallom and Shoham, 2003, Microbial hemicellulases, *Current Opinion In Microbiology* 6(3): 219-228). Hemicellulases are key components in the degradation of plant biomass. Examples of hemicellulases include, but are not limited to, an acetylmannan

esterase, an acetylxylan esterase, an arabinanase, an arabinofuranosidase, a coumaric acid esterase, a feruloyl esterase, a galactosidase, a glucuronidase, a glucuronoyl esterase, a mannanase, a mannosidase, a xylanase, and a xylosidase. The substrates of these enzymes, the hemicelluloses, are a heterogeneous group of branched and linear polysaccharides that are bound via hydrogen bonds to the cellulose microfibrils in the plant cell wall, crosslinking them into a robust network. Hemicelluloses are also covalently attached to lignin, forming together with cellulose a highly complex structure. The variable structure and organization of hemicelluloses require the concerted action of many enzymes for its complete degradation. The catalytic modules of hemicellulases are either glycoside hydrolases (GHs) that hydrolyze glycosidic bonds, or carbohydrate esterases (CEs), which hydrolyze ester linkages of acetate or ferulic acid side groups. These catalytic modules, based on homology of their primary sequence, can be assigned into GH and CE families. Some families, with an overall similar fold, can be further grouped into clans, marked alphabetically (e.g., GH-A). A most informative and updated classification of these and other carbohydrate active enzymes is available in the Carbohydrate-Active Enzymes (CAZy) database. Hemicellulolytic enzyme activities can be measured according to Ghose and Bisaria, 1987, Pure & Appl. Chem. 59: 1739-1752, at a suitable temperature, e.g., 50°C, 55°C, or 60°C, and pH, e.g., 5.0 or 5.5.

Alpha-Amylases (alpha-1,4-glucan-4-glucanohydrolases, EC 3.2.1.1) are a group of enzymes, which catalyze the hydrolysis of starch and other linear and branched 1,4-glucosidic oligo- and polysaccharides.

Glucoamylases (glucan 1,4-alpha-glucosidase, EC 3.2.1.3) are a group of enzymes, which catalyze the hydrolysis of terminal (1→4)-linked α-D-glucose residues successively from non-reducing ends of the chains with release of beta-D-glucose.

Detailed Description of the Invention

Retrograded starch or resistant starch (RS) potentially forms during the heating-cooling cycles in typical conventional corn to ethanol processes. These starch structures are thermally very stable and do not break down in normal cooking operations and temperatures. If present this form of starch can therefore not be degraded by the normal amylolytic enzymes used during liquefaction and saccharification, such as e.g., alpha-amylase and glucoamylase. It is the purpose of the present invention to provide methods and compositions useful for degrading retrograded starch. The retrograded starch may in one embodiment be formed when a starch containing material is applied in a starch conversion process. This could be any process involving heating cooling cycles. One particular example could be, e.g., in conventional starch to ethanol processes in which a liquefaction step is employed. By providing efficient degradation of retrograded starch material, end product yield may be improved.

In WO 2010/059413 a polypeptide isolated from *Aspergillus nidulans*, termed X143, was shown to have retrograded starch degrading activity. The main aspect of the present invention relates to improving the retrograded starch degrading activity of the X143 polypeptide.

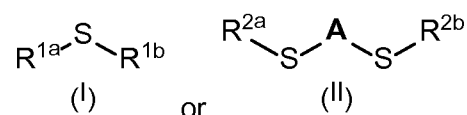
The present invention provides compositions comprising: (a) an X143 polypeptide having retrograded starch degrading activity and (b) a sulfur-containing compound, and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the X143 polypeptide having retrograded starch degrading activity and the sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, enhances degradation of retrograded starch material by the X143 polypeptide.

Furthermore, the present invention provides methods for degrading or converting a retrograded starch material, comprising: treating the retrograded starch material with an enzyme composition in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition.

Sulfur-Containing Compounds

According to the present invention, the sulfur-containing compound may be any suitable compound comprising one or more sulfur atoms. In one aspect, the sulfur-containing compound comprises a moiety selected from thionyl, thioether, sulfinyl, sulfonyl, sulfamide, sulfonamide, sulfonic acid, and sulfonic ester.

In another aspect, sulfur-containing compound is a compound of formula (I) or (II):



wherein

R^{1a} , R^{2a} , and R^{2b} are each independently hydrogen, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl;

5 R^{1b} is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl;

A is a bond or an optionally substituted moiety selected from alkylene, alkenylene, alkynylene, cycloalkylene, heterocycloalkylene, arylene, and heteroarylene;

10 wherein R^{1a} and R^{1b} together with the attached sulfur may form an optionally substituted aromatic or non-aromatic ring;

and wherein when A is a bond, one of R^{2a} and R^{2b} is not hydrogen;

or a salt or solvate thereof.

In another aspect, the sulfur-containing compound is a compound of formula (I). In another aspect, the sulfur-containing compound is a compound of formula (II).

15 In some aspects of formula (I), R^{1a} is hydrogen. In some aspects, when R^{1a} is hydrogen, the sulfur-containing compound is other than glutathione. In some aspects, R^{1a} is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R^{1a} is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, 20 cycloalkyl-alkyl. In some aspects, R^{1a} is an optionally substituted moiety selected from alkyl and cycloalkyl. In some aspects, R^{1a} is hydrogen or an optionally substituted alkyl. In some aspects, R^{1a} is an optionally substituted alkyl (e.g., an optionally substituted C_1 - C_{10} alkyl, such as an optionally substituted methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *t*-butyl, or *n*-pentyl). In some aspects, R^{1a} is an optionally substituted moiety selected from aryl, aralkyl, heteroaryl, and 25 heteroaralkyl. In some aspects, R^{1a} is an optionally substituted moiety selected from aryl and aralkyl. In some aspects, R^{1a} is an optionally substituted aryl (e.g., an optionally substituted 5 or 6-membered aryl). In some aspects, R^{1a} is an optionally substituted moiety selected from heteroaryl and heteroaralkyl. In some aspects, R^{1a} is an optionally substituted heteroaryl (e.g., an optionally substituted 5 or 6-membered heteroaryl). In some aspects, R^{1a} is an optionally 30 substituted moiety selected from phenyl, pyrazolyl, furanyl, imidazolyl, isoxazolyl, oxadiazolyl, oxazolyl, pyrrolyl, pyridyl, pyrimidyl, pyridazinyl, thiazolyl, triazolyl, thienyl, dihydrothienopyrazolyl, thianaphthenyl, carbazolyl, benzimidazolyl, benzothienyl, benzofuranyl, indolyl, quinolinyl, benzotriazolyl, benzothiazolyl, benzooxazolyl, benzimidazolyl, isoquinolinyl, isoindolyl, acridinyl, benzoisazolyl, dimethylhydantoin, pyrazinyl, tetrahydrofuranyl, pyrrolinyl, 35 pyrrolidinyl, morpholinyl, indolyl, diazepinyl, azepinyl, thiepinyl, piperidinyl, and oxepinyl. In some aspects, R^{1a} is an optionally substituted phenyl.

In some aspects of formula (I), R^{1b} is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R^{1b} is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl. In some aspects, R^{1b} is an optionally substituted moiety selected from alkyl and cycloalkyl. In some aspects, R^{1b} is an optionally substituted alkyl (e.g., an optionally substituted C_1 - C_{10} alkyl, such as an optionally substituted methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *t*-butyl, or *n*-pentyl). In some aspects, R^{1b} is an optionally substituted moiety selected from aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R^{1b} is an optionally substituted moiety selected from aryl and aralkyl. In some aspects, R^{1b} is an optionally substituted aryl (e.g., an optionally substituted 5 or 6-membered aryl). In some aspects, R^{1b} is an optionally substituted moiety selected from heteroaryl and heteroaralkyl. In some aspects, R^{1b} is an optionally substituted heteroaryl (e.g., an optionally substituted 5 or 6-membered heteroaryl). In some aspects, R^{1b} is an optionally substituted moiety selected from phenyl, pyrazolyl, furanyl, imidazolyl, isoxazolyl, oxadiazolyl, oxazolyl, pyrrolyl, pyridyl, pyrimidyl, pyridazinyl, thiazolyl, triazolyl, thienyl, dihydrothieno-pyrazolyl, thianaphthenyl, carbazolyl, benzimidazolyl, benzothienyl, benzofuranyl, indolyl, quinolyl, benzotriazolyl, benzothiazolyl, benzooxazolyl, benzimidazolyl, isoquinolyl, isoindolyl, acridinyl, benzoisazolyl, dimethylhydantoin, pyrazinyl, tetrahydrofuranyl, pyrrolinyl, pyrrolidinyl, morpholinyl, indolyl, diazepinyl, azepinyl, thiepinyl, piperidinyl, and oxepinyl. In some aspects, R^{1b} is an optionally substituted phenyl.

In some aspects of formula (I), R^{1a} and R^{1b} are identical. In some aspects, R^{1a} and R^{1b} are different. In some aspects, R^{1a} and R^{1b} together with the attached sulfur form an optionally substituted aromatic or non-aromatic ring. In some aspects, R^{1a} and R^{1b} together with the attached sulfur form an optionally substituted aromatic ring. In some aspects, R^{1a} and R^{1b} together with the attached sulfur form an optionally substituted thiophene, thiazole; benzothiophene, benzo[c]thiophene, or benzothiazole. In some aspects, R^{1a} and R^{1b} together with the attached sulfur form an optionally substituted non-aromatic ring. In some aspects, R^{1a} and R^{1b} together with the attached sulfur form an optionally substituted tetrahydrothiophene, 2,3-dihydrothiophene, or 2,5-dihydrothiophene.

In some aspects of formula (II), R^{2a} is hydrogen. In some aspects, R^{2a} is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R^{2a} is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl. In some aspects, R^{2a} is an optionally substituted moiety selected from alkyl and cycloalkyl. In some aspects, R^{2a} is hydrogen or an optionally substituted alkyl. In some aspects, R^{2a} is an optionally substituted alkyl (e.g., an optionally substituted C_1 - C_{10} alkyl, such as an optionally substituted methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *t*-butyl, or *n*-pentyl). In some

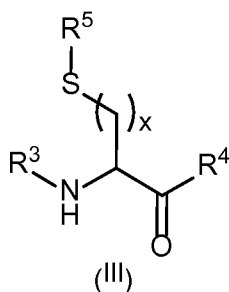
aspects, R^{2a} is an optionally substituted moiety selected from aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R^{2a} is an optionally substituted moiety selected from aryl and aralkyl. In some aspects, R^{2a} is an optionally substituted aryl (e.g., an optionally substituted 5 or 6-membered aryl). In some aspects, R^{2a} is an optionally substituted moiety selected from heteroaryl and heteroaralkyl. In some aspects, R^{2a} is an optionally substituted heteroaryl (e.g., an optionally substituted 5 or 6-membered heteroaryl). In some aspects, R^{2a} is an optionally substituted moiety selected from phenyl, pyrazolyl, furanyl, imidazolyl, isoxazolyl, oxadiazolyl, oxazolyl, pyrrolyl, pyridyl, pyrimidyl, pyridazinyl, thiazolyl, triazolyl, thienyl, dihydrothienopyrazolyl, thianaphthenyl, carbazolyl, benzimidazolyl, benzothienyl, benzofuranyl, indolyl, quinolyl, benzotriazolyl, benzothiazolyl, benzooxazolyl, benzimidazolyl, isoquinolyl, isoindolyl, acridinyl, benzoisazolyl, dimethylhydantoin, pyrazinyl, tetrahydrofuranyl, pyrrolinyl, pyrrolidinyl, morpholinyl, indolyl, diazepinyl, azepinyl, thiepinyl, piperidinyl, and oxepinyl. In some aspects, R^{2a} is an optionally substituted phenyl.

In some aspects of formula (II), R^{2b} is hydrogen. In some aspects, R^{2b} is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R^{2b} is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl. In some aspects, R^{2b} is an optionally substituted moiety selected from alkyl and cycloalkyl. In some aspects, R^{2b} is hydrogen or an optionally substituted alkyl. In some aspects, R^{2b} is an optionally substituted alkyl (e.g., an optionally substituted C_1 - C_{10} alkyl, such as an optionally substituted methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *t*-butyl, or *n*-pentyl). In some aspects, R^{2b} is an optionally substituted moiety selected from aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R^{2b} is an optionally substituted moiety selected from aryl and aralkyl. In some aspects, R^{2b} is an optionally substituted aryl (e.g., an optionally substituted 5 or 6-membered aryl). In some aspects, R^{2b} is an optionally substituted moiety selected from heteroaryl and heteroaralkyl. In some aspects, R^{2b} is an optionally substituted heteroaryl (e.g., an optionally substituted 5 or 6-membered heteroaryl). In some aspects, R^{2b} is an optionally substituted moiety selected from phenyl, pyrazolyl, furanyl, imidazolyl, isoxazolyl, oxadiazolyl, oxazolyl, pyrrolyl, pyridyl, pyrimidyl, pyridazinyl, thiazolyl, triazolyl, thienyl, dihydrothienopyrazolyl, thianaphthenyl, carbazolyl, benzimidazolyl, benzothienyl, benzofuranyl, indolyl, quinolyl, benzotriazolyl, benzothiazolyl, benzooxazolyl, benzimidazolyl, isoquinolyl, isoindolyl, acridinyl, benzoisazolyl, dimethylhydantoin, pyrazinyl, tetrahydrofuranyl, pyrrolinyl, pyrrolidinyl, morpholinyl, indolyl, diazepinyl, azepinyl, thiepinyl, piperidinyl, and oxepinyl. In some aspects, R^{2b} is an optionally substituted phenyl.

In some aspects of formula (II), R^{2a} and R^{2b} are identical. In some aspects, R^{2a} and R^{2b} are different.

In some aspects of formula (II), A is a bond. In some aspects, A is an optionally substituted moiety selected from alkylene, alkenylene, alkynylene, cycloalkylene, heterocycloalkylene, arylene, and heteroarylene. In some aspects, A is an optionally substituted moiety selected from alkylene, alkenylene, alkynylene, and cycloalkylene. In some aspects, A is an optionally substituted moiety selected from alkylene and cycloalkylene. In some aspects, A is an optionally substituted alkylene (e.g., an optionally substituted C₁-C₁₀ alkylene, such as an optionally substituted methylene, ethylene, *n*-propylene, isopropylene, *n*-butylene, *t*-butylene, or *n*-pentylene). In some aspects, A is an optionally substituted moiety selected from arylene and heteroarylene. In some aspects, A is an optionally substituted arylene (e.g., an optionally substituted 5 or 6-membered arylene). In some aspects, A is an optionally substituted heteroarylene (e.g., an optionally substituted 5 or 6-membered heteroarylene). In some aspects, A is an optionally substituted moiety selected from phenylene, pyrazolylen, furanylen, imidazolylen, isoxazolylen, oxadiazolylen, oxazolylen, pyrrolylen, pyridylen, pyrimidylen, pyridazinylen, thiazolylen, triazolylen, thienylen, dihydrothieno-pyrazolylen, thianaphthenylen, carbazolylen, benzimidazolylen, benzothienylen, benzofuranylen, indolylen, quinolinylene, benzotriazolylen, benzothiazolylen, benzooxazolylen, benzimidazolylen, isoquinolinylene, isoindolylen, acridinylen, benzoisazolylen, dimethylhydantoinene, pyrazinylen, tetrahydrofuranylen, pyrrolinylen, pyrrolidinylene, morpholinylene, indolylen, diazepinylen, azepinylen, thiepinylene, piperidinylene, and oxepinylen. In some aspects, A is an optionally substituted phenylene.

In another aspect, the sulfur-containing compound is a compound of formula (III):



wherein

x is 1 or 2;

R³ is hydrogen, -C(O)R⁶, -C(O)OR⁷, -C(O)NHR⁸, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl;

R⁴ is hydrogen, -OR⁹, -N(R¹⁰)(R¹¹), or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl;

R⁵ is hydrogen, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl;

R⁶, R⁷, R⁸, R⁹, R¹⁰, and R¹¹ are each independently hydrogen, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl;

or a salt or solvate thereof.

In some aspects of formula (III), x is 1. In some aspects, x is 2.

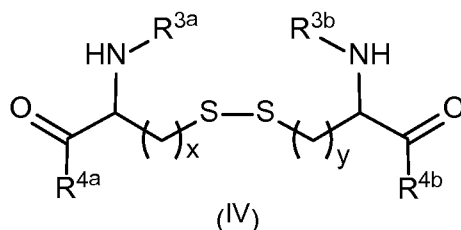
In some aspects of formula (III), R³ is hydrogen, -C(O)R⁶, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R³ is hydrogen, -C(O)R⁶, or an optionally substituted amino acid moiety; and wherein R⁶ is hydrogen or an optionally substituted alkyl. In some aspects, R³ is hydrogen, or an optionally substituted amino acid moiety. In some aspects, R³ is hydrogen. In some aspects, R³ is an optionally substituted amino acid moiety. In some aspects, when R⁵ is hydrogen, the sulfur-containing compound is other than glutathione.

In some aspects of formula (III), R⁴ is -OR⁹, -N(R¹⁰)(R¹¹), or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl; wherein R⁹, R¹⁰, and R¹¹ are each independently hydrogen, or an optionally substituted alkyl or cycloalkyl. In some aspects, R⁴ is -OR⁹, -N(R¹⁰)(R¹¹), an optionally substituted moiety selected from alkyl, or an optionally substituted amino acid moiety; wherein R⁹, R¹⁰, and R¹¹ are each independently hydrogen, or an optionally substituted alkyl or cycloalkyl. In some aspects, R⁴ is -OH or an optionally substituted amino acid moiety. In some aspects, R⁴ is -OH. In some aspects, R⁴ is an optionally substituted amino acid moiety.

In some aspects of formula (III), R⁵ is hydrogen. In some aspects, R⁵ is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R⁵ is hydrogen, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl. In some aspects, R⁵ is hydrogen, or an optionally substituted moiety selected from alkyl and cycloalkyl. In some aspects, R⁵ is hydrogen or an optionally substituted alkyl. In some aspects, R⁵ is an optionally substituted alkyl (e.g., an optionally substituted C₁-C₁₀ alkyl, such as an optionally substituted methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *t*-butyl, or *n*-pentyl). In some aspects, R⁵ is an optionally substituted moiety selected from aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R⁵ is an optionally substituted moiety selected from aryl and aralkyl. In some aspects, R⁵ is an optionally substituted aryl (e.g., an optionally substituted 5 or 6-membered aryl). In some aspects, R⁵ is an optionally substituted moiety

selected from heteroaryl and heteroaralkyl. In some aspects, R⁵ is an optionally substituted heteroaryl (e.g., an optionally substituted 5 or 6-membered heteroaryl). In some aspects, R⁵ is an optionally substituted moiety selected from phenyl, pyrazolyl, furanyl, imidazolyl, isoxazolyl, oxadiazolyl, oxazolyl, pyrrolyl, pyridyl, pyrimidyl, pyridazinyl, thiazolyl, triazolyl, thienyl, dihydrothieno-pyrazolyl, thianaphthenyl, carbazolyl, benzimidazolyl, benzothienyl, benzofuranyl, indolyl, quinoliny, benzotriazolyl, benzothiazolyl, benzooxazolyl, benzimidazolyl, isoquinoliny, isoindolyl, acridinyl, benzoisazolyl, dimethylhydantoin, pyrazinyl, tetrahydrofuranyl, pyrrolinyl, pyrrolidinyl, morpholinyl, indolyl, diazepinyl, azepinyl, thiepinyl, piperidinyl, and oxepinyl. In some aspects, R⁵ is an optionally substituted phenyl.

In another aspect, the sulfur-containing compound is a compound of formula (IV):



wherein

x and y are each independently 1 or 2;

R^{3a} and R^{3b} are each independently hydrogen, -C(O)R⁶, -C(O)OR⁷, -C(O)NHR⁸, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl;

R^{4a} and R^{4b} are each independently hydrogen, -OR⁹, -N(R¹⁰)(R¹¹), or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl;

R⁶, R⁷, R⁸, R⁹, R¹⁰, and R¹¹ are each independently hydrogen, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl;

or a salt or solvate thereof.

In some aspects of formula (IV), x is 1. In some aspects, x is 2.

In some aspects of formula (IV), y is 1. In some aspects, y is 2.

In some aspects of formula (IV), R^{3a} is hydrogen, -C(O)R⁶, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R^{3a} is hydrogen, -C(O)R⁶, or an optionally substituted amino acid moiety; and wherein R⁶ is hydrogen or an optionally substituted alkyl. In some aspects, R^{3a} is hydrogen, or an optionally substituted amino acid moiety. In some aspects, R^{3a} is hydrogen. In some aspects, R^{3a} is an optionally substituted amino acid moiety.

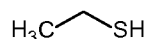
In some aspects of formula (IV), R^{3b} is hydrogen, $-C(O)R^6$, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl. In some aspects, R^{3b} is hydrogen, $-C(O)R^6$, or an optionally substituted amino acid moiety; and wherein R^6 is hydrogen or an optionally substituted alkyl.

In some aspects, R^{3b} is hydrogen, or an optionally substituted amino acid moiety. In some aspects, R^{3b} is hydrogen. In some aspects, R^{3b} is an optionally substituted amino acid moiety.

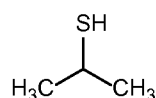
In some aspects of formula (IV), R^{4a} is $-OR^9$, $-N(R^{10})(R^{11})$, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl; wherein R^9 , R^{10} , and R^{11} are each independently hydrogen, or an optionally substituted alkyl or cycloalkyl. In some aspects, R^{4a} is $-OR^9$, $-N(R^{10})(R^{11})$, an optionally substituted moiety selected from alkyl, or an optionally substituted amino acid moiety; wherein R^9 , R^{10} , and R^{11} are each independently hydrogen, or an optionally substituted alkyl or cycloalkyl. In some aspects, R^{4a} is $-OH$ or an optionally substituted amino acid moiety. In some aspects, R^{4a} is $-OH$. In some aspects, R^{4a} is an optionally substituted amino acid moiety.

In some aspects of formula (IV), R^{4b} is $-OR^9$, $-N(R^{10})(R^{11})$, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl; wherein R^9 , R^{10} , and R^{11} are each independently hydrogen, or an optionally substituted alkyl or cycloalkyl. In some aspects, R^{4b} is $-OR^9$, $-N(R^{10})(R^{11})$, an optionally substituted moiety selected from alkyl, or an optionally substituted amino acid moiety; wherein R^9 , R^{10} , and R^{11} are each independently hydrogen, or an optionally substituted alkyl or cycloalkyl. In some aspects, R^{4b} is $-OH$ or an optionally substituted amino acid moiety. In some aspects, R^{4b} is $-OH$. In some aspects, R^{4b} is an optionally substituted amino acid moiety.

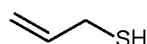
In another aspect, the sulfur-containing compound is selected from the group consisting of:



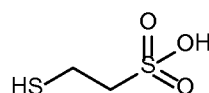
(I-1): Ethanethiol;



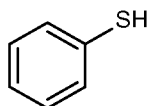
(I-2): 2-Propanethiol;



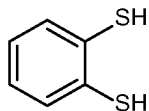
(I-3): 2-Propene-1-thiol;



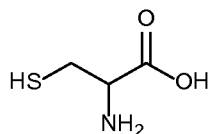
(I-4): 2-Mercaptoethanesulfonic acid (coenzyme M);



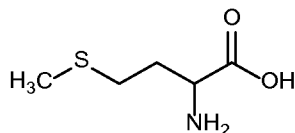
(I-6): Benzenethiol or thiophenol;



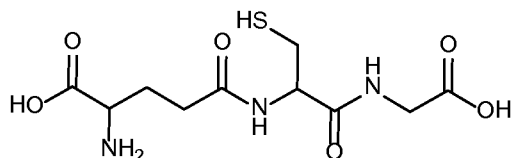
(I-7): Benzene-1,2-dithiol;



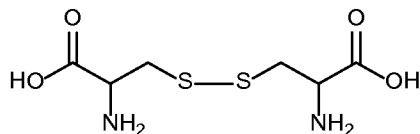
(I-5): Cysteine;



(II-1): Methionine;



(II-2): Glutathione;



(II-3): Cystine;

or a salt or solvate thereof.

In some aspects of the sulfur-containing compound described herein (e.g., a compound of formula I, II, III, or IV), the compound is other than glutathione. In another aspect, the sulfur-containing compound is L-cysteine. In another aspect, the sulfur-containing compound is glutathione. In another aspect, the sulfur-containing compound is cystine.

In some aspects, the sulfur-containing compound described herein (e.g., a compound of formula I, II, III, or IV) is in substantially pure form. With respect to the sulfur-containing compounds, unless otherwise stated, "substantially pure" intends a preparation of the sulfur-containing compound that contains no more than 15% impurity, wherein the impurity intends compounds other than the sulfur-containing compound, but does not include other forms of the sulfur-containing compound (e.g., different salt form or a different stereoisomer, conformer, rotamer, or tautomer of the analog depicted). In one variation, a preparation of substantially pure sulfur-containing compound is provided wherein the preparation contains no more than 25% impurity, or no more than 20% impurity, or no more than 10% impurity, or no more than 5% impurity, or no more than 3% impurity, or no more than 1% impurity, or no more than 0.5% impurity.

In some aspects the sulfur-containing compound described herein (e.g., a compound of formula I, II, III, or IV) is not in substantially pure form. For example, the sulfur-containing compound may be added or supplemented as part of an impure composition (e.g., unpurified biological material) wherein the composition is rich in the compound or one or more (e.g., several) chemical precursors thereof. In a one aspect, an impure composition (e.g., unpurified biological material) comprising one or more (e.g., several) sulfur-containing compounds is pretreated, e.g., as described herein for cellulosic material, and/or added to cellulosic material and/or combined with the cellulosic material prior to pretreatment of the cellulosic material. In another aspect, an impure composition (e.g., unpurified biological material) comprising one or more (e.g., several) sulfur-containing compounds is added to an enzyme composition involved in saccharification, enhancement of saccharification, liquefaction, etc. In another aspect, an impure composition (e.g., unpurified biological material) comprising one or more (e.g., several) sulfur-containing compounds is added to a fermentation or simultaneous saccharification-fermentation reaction. In any of these aspects, the impure composition comprising a sulfur-containing compound (e.g., unpurified biological material) is a preparation that contains more than 0.5% impurity, or more than 1% impurity, or more than 3% impurity, or more than 5% impurity, or more than 10% impurity, or more than 20% impurity, or more than 30% impurity, or more than 40% impurity, or more than 50% impurity, or more than 60% impurity, or more than 70% impurity, or more than 80% impurity, or more than 90% impurity, or more than 95% impurity, or more than 97% impurity, or more than 98% impurity, or more than 99% impurity.

The sulfur-containing compounds described herein (e.g., a compound of formula I, II, III, or IV) and methods of using the same, unless otherwise stated, include all solvate and/or hydrate forms. In some aspects, the sulfur-containing compounds described herein can exist in unsolvated forms as well as solvated forms (i.e., solvates). The sulfur-containing compounds may also include hydrated forms (i.e., hydrates).

The sulfur-containing compounds described herein (e.g., a compound of formula I, II, III, or IV), as well as methods of using such compounds, unless otherwise stated, include all salt forms of the compounds. The compounds also include all non-salt forms of any salt of a sulfur-containing compound described herein, as well as other salts of any salt of a sulfur-containing compound described herein. The desired salt of a basic functional group of a sulfur-containing compound may be prepared by methods known to those of skill in the art by treating the compound with an acid. The desired salt of an acidic functional group of a sulfur-containing compound can be prepared by methods known to those of skill in the art by treating the compound with a base. Examples of inorganic salts of acid compounds include, but are not limited to, alkali metal and alkaline earth salts, such as sodium salts, potassium salts, magnesium salts, bismuth salts, and calcium salts; ammonium salts; and aluminum salts. Examples of organic salts of acid compounds include, but are not limited to, procaine,

dibenzylamine, N-ethylpiperidine, N,N'-dibenzylethylenediamine, trimethylamine, and triethylamine salts. Examples of inorganic salts of base compounds include, but are not limited to, hydrochloride and hydrobromide salts. Examples of organic salts of base compounds include, but are not limited to, tartrate, citrate, maleate, fumarate, and succinate.

5 Unless stereochemistry is explicitly indicated in a chemical structure or chemical name, the chemical structure or chemical name is intended to embrace all possible stereoisomers, conformers, rotamers, and tautomers of the sulfur-containing compounds depicted. For example, a sulfur-containing compound containing a chiral carbon atom is intended to embrace both the (*R*) enantiomer and the (*S*) enantiomer, as well as mixtures of enantiomers, including
10 racemic mixtures; and a sulfur-containing compound containing two chiral carbons is intended to embrace all enantiomers and diastereomers (including (*R,R*), (*S,S*), (*R,S*), and (*R,S*) isomers). In some aspects, a sulfur-containing compound described herein (e.g., a compound of formula I, II, III, or IV) is in the form of the (*R*) enantiomer. In some aspects, a sulfur-containing compound described herein (e.g., a compound of formula I, II, III, or IV) is in the form
15 of the (*S*) enantiomer. Included in all uses of the sulfur-containing compounds disclosed herein, is any or all of the stereochemical, enantiomeric, diastereomeric, conformational, rotameric, tautomeric, solvate, hydrate, and salt forms of the compounds as described.

In another particular embodiment the sulfur-containing compound is a reducing agent and is selected from sulfur oxyanion, sulfur oxides, sulfhydryl reagent, sulfur (S), and sulfide. In
20 one preferable embodiment, Sulfur oxyanion is selected from sulfur(IV) oxyanion, sulfur(III) oxyanion, sulfur(II) oxyanion, and thiosulfate ($S_2O_3^{2-}$). In a more preferable embodiment, Sulfur(IV) oxyanion is selected from metabisulfite, for example, potassium metabisulfite ($K_2S_2O_5$), and sodium metabisulfite ($Na_2S_2O_5$); bisulfite, for example, sodium bisulfite ($NaHSO_3$); sulphite, for example, sodium sulfite (Na_2SO_3). In a particularly preferred
25 embodiment the Sulfur(IV) oxyanion is selected from sodium bisulfite ($NaHSO_3$).

The effective amount of the sulfur-containing compound can depend on one or more (e.g., several) factors including, but not limited to, the mixture of component amylolytic enzymes, the starch substrate, the concentration of starch substrate, the pretreatment(s) of the starch substrate, , temperature, and reaction time.

30 The sulfur-containing compound is preferably present in an amount that is not limiting with regard to the X143 polypeptide having retrograded starch degrading activity, amylolytic enzyme(s), and starch. In one aspect, one or more (e.g., several) sulfur-containing compounds are used in any of the methods of the present invention.

In one aspect, an effective amount of the sulfur-containing compound is about 0.1 μ M to
35 about 1 M, e.g., about 0.5 μ M to about 0.75 M, about 0.75 μ M to about 0.5 M, about 1 μ M to about 0.25 M, about 1 μ M to about 0.1 M, about 5 μ M to about 50 mM, about 10 μ M to about 25 mM, about 50 μ M to about 25 mM, about 10 μ M to about 10 mM, about 5 μ M to about 5 mM, or

about 0.1 mM to about 1 mM. In another aspect, an effective amount of the sulfur-containing compound is about 0.1 μ M to about 1 M. In another aspect, an effective amount of the sulfur-containing compound is about 0.5 μ M to about 0.75 M. In another aspect, an effective amount of the sulfur-containing compound is about 0.75 μ M to about 0.5 M. In another aspect, an effective amount of the sulfur-containing compound is about 1 μ M to about 0.25 M. In another aspect, an effective amount of the sulfur-containing compound is about 1 μ M to about 0.1 M. In another aspect, an effective amount of the sulfur-containing compound is about 5 μ M to about 50 mM. In another aspect, an effective amount of the sulfur-containing compound is about 10 μ M to about 25 mM. In another aspect, an effective amount of the sulfur-containing compound is about 50 μ M to about 25 mM. In another aspect, an effective amount of the sulfur-containing compound is about 10 μ M to about 10 mM. In another aspect, an effective amount of the sulfur-containing compound is about 5 μ M to about 5 mM. In another aspect, an effective amount of the sulfur-containing compound is about 0.1 mM to about 1 mM.

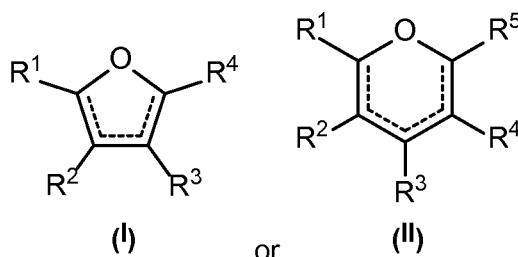
In another aspect of the present invention, the sulfur-containing compound(s) may be recycled from a completed saccharification or completed saccharification and fermentation to a new saccharification. The sulfur-containing compound(s) can be recovered using standard methods in the art, e.g., filtration/centrifugation pre- or post-distillation, to remove residual solids, cellular debris, *etc.* and then recirculated to the new saccharification.

Heterocyclic Compounds

In the methods and compositions of the present invention, the heterocyclic compound may be any suitable compound, such as an optionally substituted aromatic or non-aromatic ring comprising a heteroatom, as described herein.

In one aspect, the heterocyclic is a compound comprising an optionally substituted heterocycloalkyl moiety or an optionally substituted heteroaryl moiety. In another aspect, the optionally substituted heterocycloalkyl moiety or optionally substituted heteroaryl moiety is an optionally substituted 5-membered heterocycloalkyl or an optionally substituted 5-membered heteroaryl moiety. In another aspect, the optionally substituted heterocycloalkyl or optionally substituted heteroaryl moiety is an optionally substituted moiety selected from pyrazolyl, furanyl, imidazolyl, isoxazolyl, oxadiazolyl, oxazolyl, pyrrolyl, pyridyl, pyrimidyl, pyridazinyl, thiazolyl, triazolyl, thienyl, dihydrothieno-pyrazolyl, thianaphthenyl, carbazolyl, benzimidazolyl, benzothienyl, benzofuranyl, indolyl, quinoliny, benzotriazolyl, benzothiazolyl, benzooxazolyl, benzimidazolyl, isoquinoliny, isoindolyl, acridinyl, benzoisazolyl, dimethylhydantoin, pyrazinyl, tetrahydrofuranyl, pyrrolinyl, pyrrolidinyl, morpholinyl, indolyl, diazepinyl, azepinyl, thiepinyl, piperidinyl, and oxepinyl. In another aspect, the optionally substituted heterocycloalkyl moiety or optionally substituted heteroaryl moiety is an optionally substituted furanyl.

In another aspect, the heterocyclic compound is a compound is of formula (I) or (II):



wherein

each bond indicated with a dashed line is single or double;

5 R^1 , R^2 , R^3 , R^4 , and R^5 are independently hydrogen, halogen, =O, -OH, -OR⁸, -CN, -NO₂, -N(R⁹)(R¹⁰), -C(O)R²⁰, -C(O)OR⁶, -C(O)NHR⁷, -OC(O)R¹¹, -NHC(O)R¹², -OC(O)OR¹³, -NHC(O)OR¹⁴, -OC(O)NHR¹⁵, -NHC(O)NHR¹⁶, -SO₂R¹⁷, -SO₂N(R¹⁸)(R¹⁹), -SR²¹, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl;

10 R^6 , R^7 , R^8 , R^9 , R^{10} , R^{11} , R^{12} , R^{13} , R^{14} , R^{15} , R^{16} , R^{18} , R^{19} , R^{20} , and R^{21} are independently hydrogen, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl; and

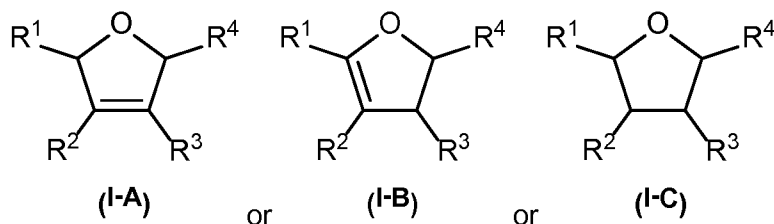
15 R^{17} is an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl; and

wherein each pair of R^1 and R^2 , R^2 and R^3 , R^3 and R^4 , and R^4 and R^5 may combine to form an optionally substituted fused ring;

or a salt or solvate thereof.

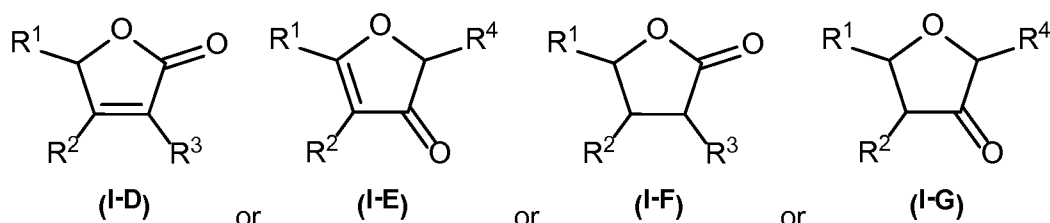
20 In another aspect of formula (I) or (II), at least one bond indicated with a dashed line is double. In another aspect, only one bond indicated with a dashed line is double.

In another aspect, the heterocyclic compound is a compound is of formula (I-A), (II-B), or (II-C):



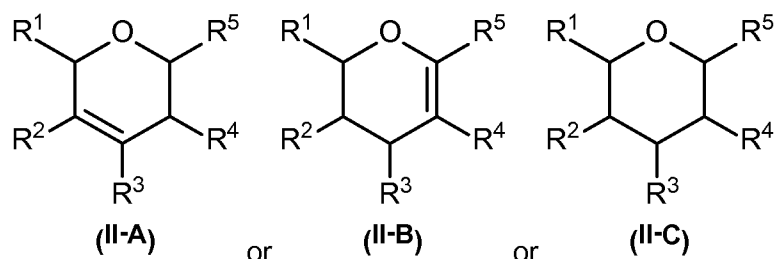
25 wherein R^1 , R^2 , R^3 , and R^4 are as defined above; or a salt or solvate thereof.

In another aspect, the heterocyclic compound is a compound is of formula (I-D), (I-E), (I-F), or (II-G):



wherein R¹, R², R³, and R⁴ are as defined above; or a salt or solvate thereof.

In another aspect, the heterocyclic compound is a compound of formula (II-A), (II-B), or (II-C):



wherein R¹, R², R³, R⁴ and R⁵ are as defined above; or a salt or solvate thereof.

In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); R¹, R², R³, R⁴, and R⁵ are independently hydrogen, halogen, =O, -OH, -OR⁸, or an optionally substituted moiety selected from alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkyl-alkyl, heterocycloalkyl, heterocycloalkyl-alkyl, aryl, aralkyl, heteroaryl, and heteroaralkyl; wherein each pair of R¹ and R², R² and R³, R³ and R⁴, and R⁴ and R⁵ may combine to form an optionally substituted fused ring. In another aspect, R¹, R², R³, R⁴, and R⁵ are independently hydrogen, halogen, =O, -OH, -OR⁸, or an optionally substituted alkyl; wherein each pair of R¹ and R², R² and R³, R³ and R⁴, and R⁴ and R⁵ may combine to form an optionally substituted fused ring. In another aspect, R¹, R², R³, R⁴, and R⁵ are independently hydrogen, =O, -OH, an optionally substituted -O-(C₁-C₁₀)alkyl, or an optionally substituted -(C₁-C₁₀)alkyl.

In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); at least one of R¹, R², R³, R⁴, and R⁵ is hydrogen. In another aspect, at least two of R¹, R², R³, R⁴, and R⁵ are hydrogen. In another aspect, at least three of R¹, R², R³, R⁴, and R⁵ are hydrogen.

In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); at least one of R¹, R², R³, R⁴, and R⁵, is an optionally substituted alkyl (e.g., an optionally substituted C₁-C₁₀ alkyl, such as an optionally substituted methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *t*-butyl, or *n*-pentyl). In another aspect, at least two of R¹, R², R³, R⁴, and R⁵, are optionally substituted alkyl.

In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); at least one of R¹, R², R³, R⁴, and R⁵ is =O. In another aspect, only one of R¹, R², R³, R⁴, and R⁵ is =O. In another aspect, R¹ is =O. In another aspect, R² is =O. In another aspect, R³ is =O. In another aspect, R⁴ is =O. In another aspect, R⁵ is =O.

In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); at least two of R^1 , R^2 , R^3 , R^4 , and R^5 are =O. In another aspect, only two of R^1 , R^2 , R^3 , R^4 , and R^5 are =O. In another aspect, R^1 and R^2 are =O. In another aspect, R^1 and R^3 are =O. In another aspect, R^1 and R^4 are =O. In another aspect, R^1 and R^5 are =O. In another aspect, R^2 and R^3 are =O. In another aspect, R^2 and R^4 are =O. In another aspect, R^2 and R^5 are =O. In another aspect, R^3 and R^4 are =O. In another aspect, R^3 and R^5 are =O. In another aspect, R^4 and R^5 are =O.

In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); three of R^1 , R^2 , R^3 , R^4 , and R^5 are =O.

In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); at least one of R^1 , R^2 , R^3 , R^4 , and R^5 is -OH. In another aspect, only one of R^1 , R^2 , R^3 , R^4 , and R^5 is -OH. In another aspect, R^1 is -OH. In another aspect, R^2 is -OH. In another aspect, R^3 is -OH. In another aspect, R^4 is -OH. In another aspect, R^5 is -OH.

In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); at least two of R^1 , R^2 , R^3 , R^4 , and R^5 are -OH. In another aspect, only two of R^1 , R^2 , R^3 , R^4 , and R^5 are -OH. In another aspect, R^1 and R^2 are -OH. In another aspect, R^1 and R^3 are -OH. In another aspect, R^1 and R^4 are -OH. In another aspect, R^1 and R^5 are -OH. In another aspect, R^2 and R^3 are -OH. In another aspect, R^2 and R^4 are -OH. In another aspect, R^2 and R^5 are -OH. In another aspect, R^3 and R^4 are -OH. In another aspect, R^3 and R^5 are -OH. In another aspect, R^4 and R^5 are -OH.

In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); at least three of R^1 , R^2 , R^3 , R^4 , and R^5 are -OH.

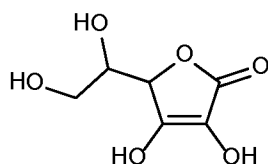
In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); at least one of R^1 , R^2 , R^3 , R^4 , and R^5 is -OH and at least one of R^1 , R^2 , R^3 , R^4 , and R^5 is =O.

In another aspect of formula (I), (I-A), (I-B), (I-C), (I-D), (I-E), (I-F), (I-G), (II), (II-A), (II-B), or (II-C); at least one pair of R^1 and R^2 , R^2 and R^3 , R^3 and R^4 , and R^4 and R^5 combine to form an optionally substituted fused ring. In another aspect, R^1 and R^2 combine to form an optionally substituted fused ring. In another aspect, R^1 and R^2 combine to form an optionally substituted fused cycloalkylene ring. In another aspect, R^1 and R^2 combine to form an optionally substituted fused arylene ring. In another aspect, R^1 and R^2 combine to form an optionally substituted fused heteroarylene ring. In another aspect, R^2 and R^3 combine to form an optionally substituted fused ring. In another aspect, R^2 and R^3 combine to form an optionally substituted fused cycloalkylene ring. In another aspect, R^2 and R^3 combine to form an optionally substituted fused arylene ring. In another aspect, R^2 and R^3 combine to form an optionally substituted fused heteroarylene ring. In another aspect, R^3 and R^4 combine to form an optionally substituted fused ring. In another aspect, R^3 and R^4 combine to form an optionally substituted fused

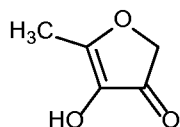
cycloalkylene ring. In another aspect, R^3 and R^4 combine to form an optionally substituted fused arylene ring. In another aspect, R^3 and R^4 combine to form an optionally substituted fused heteroarylene ring. In another aspect, R^4 and R^5 combine to form an optionally substituted fused ring. In another aspect, R^4 and R^5 combine to form an optionally substituted fused cycloalkylene ring. In another aspect, R^4 and R^5 combine to form an optionally substituted fused arylene ring. In another aspect, R^4 and R^5 combine to form an optionally substituted fused heteroarylene ring.

In another aspect, only one pair of R^1 and R^2 , R^2 and R^3 , R^3 and R^4 , and R^4 and R^5 combine to form an optionally substituted fused ring.

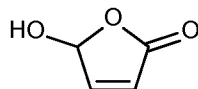
In another aspect, the heterocyclic compound is selected from the group consisting of:



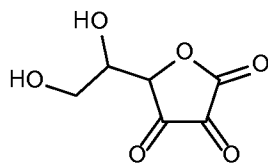
(I-1): (1,2-Dihydroxyethyl)-3,4-dihydroxyfuran-2(5H)-one;



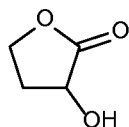
(I-2): 4-Hydroxy-5-methyl-3-furanone;



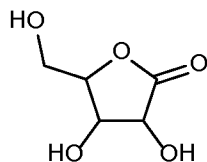
(I-3): 5-Hydroxy-2(5H)-furanone;



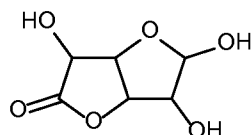
(I-4): [1,2-Dihydroxyethyl]furan-2,3,4(5H)-trione;



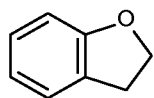
(I-5): α-hydroxy-γ-butyrolactone;



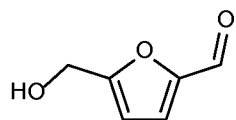
(I-6): Ribonic γ-lactone;



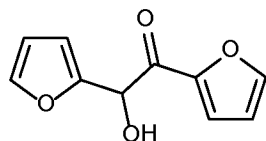
(I-7): Glucuronic acid γ-lactone;



(I-8): Dihydrobenzofuran;

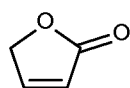


(I-9): 5-(hydroxymethyl)furfural;

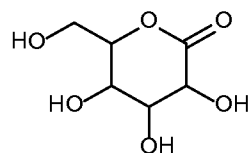


5

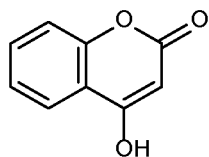
(I-10): Furoin;



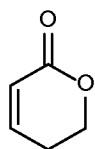
(I-11): 2(5H)-Furanone;



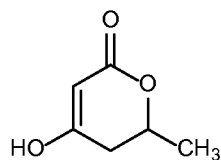
10

(II-1): Gluconic acid δ -lactone;

(II-2): 4-Hydroxycoumarin;

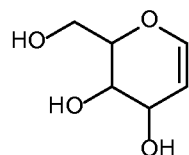


(II-3): 5,6-Dihydro-2H-pyran-2-one;

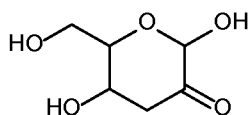


15

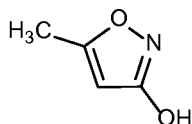
(II-4): 5,6-Dihydro-4-hydroxy-6-methyl-2H-pyran-2-one;



(II-5): 1,5-anhydro-2-deoxy-arabino-hex-1-enitol;



(II-6): 3-deoxy-erythro-hexosulose;



3-Hydroxy-5-methylisoxazole;

5 or a salt or solvate thereof.

In some aspects, the heterocyclic compound described herein (e.g., a compound of formula I, I-A, I-B, I-C, I-D, I-E, I-F, I-G, II, II-A, II-B, or II-C) is in substantially pure form. With respect to the heterocyclic compounds, unless otherwise stated, "substantially pure" intends a preparation of the heterocyclic compound that contains no more than 15% impurity, wherein the impurity intends compounds other than the heterocyclic compound, but does not include other forms of the heterocyclic compound (e.g., different salt form or a different stereoisomer, conformer, rotamer, or tautomer of the analog depicted). In one variation, a preparation of substantially pure heterocyclic compound is provided wherein the preparation contains no more than 25% impurity, or no more than 20% impurity, or no more than 10% impurity, or no more than 5% impurity, or no more than 3% impurity, or no more than 1% impurity, or no more than 0.5% impurity.

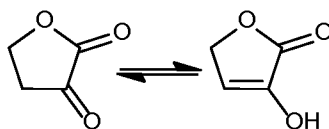
In some aspects the heterocyclic compound described herein (e.g., a compound of formula I, I-A, I-B, I-C, I-D, I-E, I-F, I-G, II, II-A, II-B, or II-C) is not in substantially pure form. For example, the heterocyclic compound may be added or supplemented as part of an impure composition (e.g., unpurified biological material) wherein the composition is rich in the compound or one or more (e.g., several) chemical precursors thereof. In a one aspect, an impure composition (e.g., unpurified biological material) comprising one or more (e.g., several) heterocyclic compounds is pretreated, e.g., as described herein for cellulosic material, and/or added to cellulosic material and/or combined with the cellulosic material prior to pretreatment of the cellulosic material. In another aspect, an impure composition (e.g., unpurified biological material) comprising one or more (e.g., several) heterocyclic compounds is added to an enzyme composition involved in saccharification, enhancement of saccharification, liquefaction, etc. In another aspect, an impure composition (e.g., unpurified biological material) comprising one or more (e.g., several) heterocyclic compounds is added to a fermentation or simultaneous saccharification-fermentation reaction. In any of these aspects, the impure composition comprising a heterocyclic compound (e.g., unpurified biological material) is a preparation that contains more than 0.5% impurity, or more than 1% impurity, or more than 3% impurity, or more than 5% impurity, or more than 10% impurity, or more than 20% impurity, or more than 30% impurity, or more than 40% impurity, or more than 50% impurity, or more than 60% impurity, or more than 70% impurity, or more than 80% impurity, or more than 90% impurity, or more than

95% impurity, or more than 97% impurity, or more than 98% impurity, or more than 99% impurity.

The heterocyclic compounds described herein (*e.g.*, a compound of formula I, I-A, I-B, I-C, I-D, I-E, I-F, I-G, II, II-A, II-B, or II-C) and methods of using the same, unless otherwise stated, include all solvate and/or hydrate forms. In some aspects, the heterocyclic compounds described herein can exist in unsolvated forms as well as solvated forms (*i.e.*, solvates). The heterocyclic compounds may also include hydrated forms (*i.e.*, hydrates).

The heterocyclic compounds described herein (*e.g.*, a compound of formula I, I-A, I-B, I-C, I-D, I-E, I-F, I-G, II, II-A, II-B, or II-C), as well as methods of using such compounds, unless otherwise stated, include all salt forms of the compounds. The compounds also include all non-salt forms of any salt of a heterocyclic compound described herein, as well as other salts of any salt of a heterocyclic compound described herein. The desired salt of a basic functional group of a heterocyclic compound may be prepared by methods known to those of skill in the art by treating the compound with an acid. The desired salt of an acidic functional group of a heterocyclic compound can be prepared by methods known to those of skill in the art by treating the compound with a base. Examples of inorganic salts of acid compounds include, but are not limited to, alkali metal and alkaline earth salts, such as sodium salts, potassium salts, magnesium salts, bismuth salts, and calcium salts; ammonium salts; and aluminum salts. Examples of organic salts of acid compounds include, but are not limited to, procaine, dibenzylamine, N-ethylpiperidine, N,N'-dibenzylethylenediamine, trimethylamine, and triethylamine salts. Examples of inorganic salts of base compounds include, but are not limited to, hydrochloride and hydrobromide salts. Examples of organic salts of base compounds include, but are not limited to, tartrate, citrate, maleate, fumarate, and succinate.

Unless stereochemistry is explicitly indicated in a chemical structure or chemical name, the chemical structure or chemical name is intended to embrace all possible stereoisomers, conformers, rotamers, and tautomers of the heterocyclic compounds depicted. For example, a heterocyclic compound containing a chiral carbon atom is intended to embrace both the (*R*) enantiomer and the (*S*) enantiomer, as well as mixtures of enantiomers, including racemic mixtures; and a heterocyclic compound containing two chiral carbons is intended to embrace all enantiomers and diastereomers (including (*R,R*), (*S,S*), (*R,S*), and (*R,S*) isomers). In some aspects, a heterocyclic compound described herein (*e.g.*, a compound of formula I, I-A, I-B, I-C, I-D, I-E, I-F, I-G, II, II-A, II-B, or II-C) is in the form of the (*R*) enantiomer. In some aspects, a heterocyclic compound described herein (*e.g.*, a compound of formula I, I-A, I-B, I-C, I-D, I-E, I-F, I-G, II, II-A, II-B, or II-C) is in the form of the (*S*) enantiomer. The chemical structure is intended to embrace all tautomeric structures. For example, a structure such as 3-hydroxy-5*H*-furan-2-one is intended to also embrace the tautomeric form of dihydrofuran 2,3-dione:



Included in all uses of the heterocyclic compounds disclosed herein, is any or all of the stereochemical, enantiomeric, diastereomeric, conformational, rotomeric, tautomeric, solvate, hydrate, and salt forms of the compounds as described.

5 The effective amount of the heterocyclic compound can depend on one or more (e.g., several) factors including, but not limited to, the mixture of component amylolytic enzymes, the starch substrate, the concentration of starch substrate, the pretreatment(s) of the starch substrate, temperature, and reaction time.

10 The heterocyclic compound is preferably present in an amount that is not limiting with regard to the X143 polypeptide having retrograded starch degrading activity, amylolytic enzyme(s), and starch.

15 In one aspect, an effective amount of the heterocyclic compound 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. In another aspect, an effective amount of the heterocyclic compound is about 0.1 μM to about 1 M. In another aspect, an effective amount of the heterocyclic compound is about 0.5 μM to about 0.75 M. In another aspect, an effective amount of the heterocyclic compound is about 0.75 μM to about 0.5 M. In another aspect, an effective amount of the heterocyclic compound is about 1 μM to about 0.25 M. In another aspect, an effective amount of the heterocyclic compound is about 1 μM to about 0.1 M. In another aspect, an effective amount of the heterocyclic compound is about 5 μM to about 50 mM. In another aspect, an effective amount of the heterocyclic compound is about 10 μM to about 25 mM. In another aspect, an effective amount of the heterocyclic compound is about 50 μM to about 25 mM. In another aspect, an effective amount of the heterocyclic compound is about 10 μM to about 10 mM. In another aspect, an effective amount of the heterocyclic compound is about 5 μM to about 5 mM. In another aspect, an effective amount of the heterocyclic compound is about 0.1 mM to about 1 mM.

20 In another aspect of the present invention, the heterocyclic compound(s) may be recycled from a completed saccharification or completed saccharification and fermentation to a new saccharification. The heterocyclic compound(s) can be recovered using standard methods in the art, e.g., filtration/centrifugation pre- or post-distillation, to remove residual solids, cellular debris, etc. and then recirculated to the new saccharification.

Cellobiose Dehydrogenases and Polynucleotides Thereof

In the methods and compositions of the present invention, the cellobiose dehydrogenase can be any cellobiose dehydrogenase. The cellobiose dehydrogenase may be present as an enzyme activity in the enzyme composition and/or as a component of one or more protein components added to the composition.

The cellobiose dehydrogenase is preferably present in an amount that is not limiting with regard to the X143 polypeptide having retrograded starch degrading activity, amylolytic enzyme(s), and starch.

In a particular embodiment the ratio of cellobiose dehydrogenase to X143 polypeptide is in the range from 1:10 to 10:1, more particularly from 5:10 to 10:5, such as e.g., 1:1.

The cellobiose dehydrogenase may be obtained from microorganisms of any genus. For purposes of the present invention, the term "obtained from" as used herein in connection with a given source shall mean that the polypeptide encoded by a nucleotide sequence is produced by the source or by a strain in which the nucleotide sequence from the source has been inserted.

In one aspect, the polypeptide obtained from a given source is secreted extracellularly.

The cellobiose dehydrogenase 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*, or *Oceanobacillus* cellobiose dehydrogenase, or a Gram negative bacterial polypeptide such as an *E. coli*, *Pseudomonas*, *Salmonella*, *Campylobacter*, *Helicobacter*, *Flavobacterium*, *Fusobacterium*, *Ilyobacter*, *Neisseria*, or *Ureaplasma* cellobiose dehydrogenase.

In one aspect, the polypeptide is a *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, or *Bacillus thuringiensis* cellobiose dehydrogenase.

In another aspect, the polypeptide is a *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, or *Streptococcus equi* subsp. *Zooepidemicus* cellobiose dehydrogenase.

In another aspect, the polypeptide is a *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, or *Streptomyces lividans* cellobiose dehydrogenase.

The cellobiose dehydrogenase may also be a fungal polypeptide, and more preferably a yeast polypeptide such as a *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* cellobiose dehydrogenase; or more preferably a filamentous fungal polypeptide such as an *Acremonium*, *Agaricus*, *Alternaria*, *Aspergillus*, *Aureobasidium*, *Botryosphaeria*, *Ceriporiopsis*, *Chaetomidium*, *Chrysosporium*, *Claviceps*,

Cochliobolus, *Coprinopsis*, *Coptotermes*, *Corynascus*, *Cryphonectria*, *Cryptococcus*, *Diplodia*, *Exidia*, *Filibasidium*, *Fusarium*, *Gibberella*, *Holomastigotoides*, *Humicola*, *Irpex*, *Lentinula*, *Leptosphaeria*, *Magnaporthe*, *Melanocarpus*, *Meripilus*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Piromyces*, *Poitrasia*,
 5 *Pseudoplectania*, *Pseudotrichonympha*, *Rhizomucor*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trichoderma*, *Trichophaea*, *Verticillium*, *Volvariella*, or *Xylaria* cellobiose dehydrogenase.

In another aspect, the polypeptide is a *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*,
 10 *Saccharomyces norbensis*, or *Saccharomyces oviformis* cellobiose dehydrogenase.

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

Examples of other cellobiose dehydrogenases and their sources are listed in Table 1.

Table 1. Published microbial cellobiose dehydrogenase sequences

Species	Accession #	Literature Reference
<i>Humicola insolens</i>	Q9P8H5	Xu <i>et al.</i> , 2001, <i>Humicola insolens</i> cellobiose dehydrogenase: cloning, redox chemistry, and "logic gate"-like dual functionality, <i>Enz. Microb. Technol.</i> 28: 744-753
<i>Irpex lacteus</i>	Q6AW20	Nozaki <i>et al.</i> , 1999, Cloning and expression of cellobiose dehydrogenase from <i>Irpex lacteus</i> . Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
<i>Pycnoporus cinnabarinus</i>	O74253	Moukha <i>et al.</i> , 1999, Cloning and analysis of <i>Pycnoporus cinnabarinus</i> cellobiose dehydrogenase, <i>Gene</i> 234: 23-33
<i>Phanerochaete</i>	Q01738	Li <i>et al.</i> , 1996, Cloning of a cDNA encoding cellobiose dehydrogenase, a hemoflavoenzyme from

Species	Accession #	Literature Reference
<i>chrysosporium</i>		Phanerochaete chrysosporium, <i>Appl. Environ. Microbiol.</i> 62: 1329-1335
<i>Coniophora puteana</i>	Q6BDD5	Kajisa <i>et al.</i> , 2004, Characterization and molecular cloning of cellobiose dehydrogenase from the brown-rot fungus <i>Coniophora puteana</i> , <i>Biosci. Bioeng.</i> 98: 57-63
<i>Athelia rolfsii</i>	Q7Z975	Zamocky <i>et al.</i> , Phylogenetic analysis of cellobiose dehydrogenases. Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases
<i>Grifola frondosa</i>	Q8J2T4	Yoshida <i>et al.</i> , 2002, Molecular cloning and characterization of a cDNA encoding cellobiose dehydrogenase from the wood-rotting fungus <i>Grifola frondosa</i> , <i>FEMS Microbiol. Lett.</i> 217: 225-230
<i>Trametes versicolor</i>	Q875J3	Stapleton <i>et al.</i> , 2004, Molecular cloning of the cellobiose dehydrogenase gene from <i>Trametes versicolor</i> and expression in <i>Pichia pastoris</i> , <i>Enzyme Microb. Technol.</i> 34: 55-63
<i>Trametes versicolor</i>	O42729	Dumonceaux <i>et al.</i> , 1998, Cloning and sequencing of a gene encoding cellobiose dehydrogenase from <i>Trametes versicolor</i> , <i>Gene</i> 210: 211-219
<i>Aspergillus fumigatus</i>	Q4WIN9	Nierman <i>et al.</i> , 2005, Genomic sequence of the pathogenic and allergenic filamentous fungus <i>Aspergillus fumigatus</i> , <i>Nature</i> 438: 1151-1156
<i>Phanerochaete chrysosporium</i>	Q12661	Raices <i>et al.</i> , 1995, Cloning and characterization of a cDNA encoding a cellobiose dehydrogenase from the white rot fungus <i>Phanerochaete chrysosporium</i> , <i>FEBS Lett.</i> 369: 233-238
<i>Myriococcum thermophilum</i>	A9XK88	Zamocky <i>et al.</i> , 2008, Cloning, sequence analysis and heterologous expression in <i>Pichia pastoris</i> of a gene encoding a thermostable cellobiose dehydrogenase from <i>Myriococcum thermophilum</i> , <i>Protein Expr. Purif.</i> 59: 258-265
<i>Aspergillus clavatus</i>	A1CFV0	Fedorova <i>et al.</i> , Genomic islands in the pathogenic filamentous fungus <i>Aspergillus fumigatus</i> , PLoS
<i>Corynascus heterothallicus</i>	O74240	Subramaniam <i>et al.</i> , Biochemical and molecular biological characterization of cellobiose dehydrogenase from <i>Sporotrichum thermophile</i> . Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases
<i>Neosartorya fischeri</i>	A1CYG2	Fedorova <i>et al.</i> , Genomic islands in the pathogenic filamentous fungus <i>Aspergillus fumigatus</i> , PLoS
<i>Aspergillus fumigatus</i>	B0XVQ8	Fedorova <i>et al.</i> , Genomic islands in the pathogenic filamentous fungus <i>Aspergillus fumigatus</i> , PLoS
<i>Aspergillus clavatus</i>	A1C890	Fedorova <i>et al.</i> , Genomic islands in the pathogenic filamentous fungus <i>Aspergillus fumigatus</i> , PLoS
<i>Neosartorya fischeri</i>	A1DIY3	Fedorova <i>et al.</i> , Genomic islands in the pathogenic filamentous fungus <i>Aspergillus fumigatus</i> , PLoS
<i>Myriococcum thermophilum</i>	A9XK87	Zamocky <i>et al.</i> , 2008, Cloning, sequence analysis and heterologous expression in <i>Pichia pastoris</i> of a gene encoding a thermostable cellobiose dehydrogenase from <i>Myriococcum thermophilum</i> , <i>Protein Expr. Purif.</i> 59: 258-265
<i>Pyrenophora tritici-repentis</i>	B2WHI7	Birren <i>et al.</i> , The Broad Institute Genome Sequencing Platform "Genome Sequence of <i>Pyrenophora tritici-repentis</i> . Submitted (MAR-2007) to the EMBL/GenBank/DDBJ databases

Species	Accession #	Literature Reference
<i>Pyrenophora tritici-repentis</i>	B2WJX3	Birren <i>et al.</i> , The Broad Institute Genome Sequencing Platform "Genome Sequence of <i>Pyrenophora tritici-repentis</i> . Submitted (MAR-2007) to the EMBL/GenBank/DDBJ databases
<i>Aspergillus fumigatus</i>	Q4WC40	Fedorova <i>et al.</i> , Genomic islands in the pathogenic filamentous fungus <i>Aspergillus fumigatus</i> , PLoS
<i>Aspergillus niger</i>	A2QD75	Pel <i>et al.</i> , 2007, Genome sequencing and analysis of the versatile cell factory <i>Aspergillus niger</i> CBS 513.88, <i>Nat. Biotechnol.</i> 25: 221-231

In another aspect, the cellobiose dehydrogenase is a *Humicola insolens* cellobiose dehydrogenase. In another aspect, the cellobiose dehydrogenase is a *Humicola insolens* DSM 1800 cellobiose dehydrogenase, (see U.S. Patent No. 6,280,976).

5 In another aspect, the cellobiose dehydrogenase is a *Myceliophthora thermophila* cellobiose dehydrogenase. In another aspect, the cellobiose dehydrogenase is a *Myceliophthora thermophila* CBS 117.65 cellobiose dehydrogenase.

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, 10 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 15 (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

Polypeptides Having Retrograded Starch Activity and Polynucleotides thereof

A polypeptide having retrograded starch degrading activity to be applied in the methods 20 and compositions of the present invention relates to isolated polypeptides comprising amino acid sequences having a degree of sequence identity to the mature polypeptide of SEQ ID NO: 2 of preferably at least 60%, more preferably at least 65%, more preferably at least 70%, more preferably at least 75%, more preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, most preferably at least 95%, and even most preferably at least 96%, 25 at least 97%, at least 98%, or at least 99%, which have retrograded starch activity (hereinafter "homologous polypeptides"). In a preferred aspect, the homologous polypeptides comprise amino acid sequences that differ by ten amino acids, preferably by five amino acids, more preferably by four amino acids, even more preferably by three amino acids, most preferably by two amino acids, and even most preferably by one amino acid from the mature polypeptide of 30 SEQ ID NO: 2.

A polypeptide having retrograded starch degrading activity preferably comprises the amino acid sequence of SEQ ID NO: 2 or an allelic variant thereof; or a fragment thereof having amylolytic enhancing activity. In a preferred aspect, the polypeptide comprises the amino acid sequence of SEQ ID NO: 2. In another preferred aspect, the polypeptide comprises the mature polypeptide of SEQ ID NO: 2. In another preferred aspect, the polypeptide comprises amino acids 19 to 385 of SEQ ID NO: 2, or an allelic variant thereof; or a fragment thereof having retrograded starch degrading activity activity. In another preferred aspect, the polypeptide comprises amino acids 19 to 385 of SEQ ID NO: 2. In another preferred aspect, the polypeptide consists of the amino acid sequence of SEQ ID NO: 2 or an allelic variant thereof; or a fragment thereof having retrograded starch degrading activity. In another preferred aspect, the polypeptide consists of the amino acid sequence of SEQ ID NO: 2. In another preferred aspect, the polypeptide consists of the mature polypeptide of SEQ ID NO: 2. In another preferred aspect, the polypeptide consists of amino acids 19 to 385 of SEQ ID NO: 2 or an allelic variant thereof; or a fragment thereof having retrograde starch degrading activity. In another preferred aspect, the polypeptide consists of amino acids 19 to 385 of SEQ ID NO: 2.

A polypeptide having retrograded starch degrading activity is encoded by polynucleotides that hybridize under preferably very low stringency conditions, more preferably low stringency conditions, more preferably medium stringency conditions, more preferably medium-high stringency conditions, even more preferably high stringency conditions, and most preferably very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, (ii) the cDNA sequence contained in the mature polypeptide coding sequence of SEQ ID NO: 1, or (iii) a full-length complementary strand of (i) or (ii) (J. Sambrook, E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, New York).

The nucleotide sequence of SEQ ID NO: 1; or a subsequence thereof; as well as the amino acid sequence of SEQ ID NO: 2; or a fragment thereof; may be used to design nucleic acid probes to identify and clone DNA encoding polypeptides having retrograded starch degrading activity from strains of different genera or species according to methods well known in the art. In particular, such probes can be used for hybridization with the genomic or cDNA of the genus or species of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Such probes can be considerably shorter than the entire sequence, but should be at least 14, preferably at least 25, more preferably at least 35, and most preferably at least 70 nucleotides in length. It is, however, preferred that the nucleic acid probe is at least 100 nucleotides in length. For example, the nucleic acid probe may be at least 200 nucleotides, preferably at least 300 nucleotides, more preferably at least 400 nucleotides, or most preferably at least 500 nucleotides in length. Even longer probes may be used, e.g., nucleic acid probes that are preferably at least 600 nucleotides, more preferably

at least 700 nucleotides, even more preferably at least 800 nucleotides, or most preferably at least 900 nucleotides in length. Both DNA and RNA probes can be used. The probes are typically labeled for detecting the corresponding gene (for example, with ^{32}P , ^3H , ^{35}S , biotin, or avidin). Such probes are encompassed by the present invention.

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

For purposes of the present invention, hybridization indicates that the nucleotide sequence hybridizes to a labeled nucleic acid probe corresponding to the mature polypeptide
15 coding sequence of SEQ ID NO: 1; the cDNA sequence contained in the mature polypeptide coding sequence of SEQ ID NO: 1; its full-length complementary strand; or a subsequence thereof; under very low to very high stringency conditions. Molecules to which the nucleic acid probe hybridizes under these conditions can be detected using, for example, X-ray film.

Preferably, the nucleic acid probe is the mature polypeptide coding sequence of SEQ ID
20 NO: 1. In another preferred aspect, the nucleic acid probe is nucleotides 55 to 1214 of SEQ ID NO: 1. In another preferred aspect, the nucleic acid probe is a polynucleotide sequence that encodes the polypeptide of SEQ ID NO: 2, or a subsequence thereof. In another preferred aspect, the nucleic acid probe is SEQ ID NO: 1.

For long probes of at least 100 nucleotides in length, very low to very high stringency
25 conditions are defined as prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 µg/ml sheared and denatured salmon sperm DNA, and either 25% formamide for very low and low stringencies, 35% formamide for medium and medium-high stringencies, or 50% formamide for high and very high stringencies, following standard Southern blotting procedures for 12 to 24 hours optimally.

30 For long probes of at least 100 nucleotides in length, the carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS preferably at 45°C (very low stringency), more preferably at 50°C (low stringency), more preferably at 55°C (medium stringency), more preferably at 60°C (medium-high stringency), even more preferably at 65°C (high stringency), and most preferably at 70°C (very high stringency).

35 For short probes of about 15 nucleotides to about 70 nucleotides in length, stringency conditions are defined as prehybridization, hybridization, and washing post-hybridization at about 5°C to about 10°C below the calculated T_m using the calculation according to Bolton and

McCarthy (1962, *Proceedings of the National Academy of Sciences USA* 48:1390) in 0.9 M NaCl, 0.09 M Tris-HCl pH 7.6, 6 mM EDTA, 0.5% NP-40, 1X Denhardt's solution, 1 mM sodium pyrophosphate, 1 mM sodium monobasic phosphate, 0.1 mM ATP, and 0.2 mg of yeast RNA per ml following standard Southern blotting procedures for 12 to 24 hours optimally.

5 For short probes of about 15 nucleotides to about 70 nucleotides in length, the carrier material is washed once in 6X SCC plus 0.1% SDS for 15 minutes and twice each for 15 minutes using 6X SSC at 5°C to 10°C below the calculated T_m .

Isolated polypeptides having retrograded starch degrading activity is encoded by polynucleotides comprising or consisting of nucleotide sequences having a degree of sequence
10 identity to the mature polypeptide coding sequence of SEQ ID NO: 1 of preferably at least 60%, more preferably at least 65%, more preferably at least 70%, more preferably at least 75%, more preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, most preferably at least 95%, and even most preferably at least 96%, at least 97%, at least 98%, or at least 99%, which encode a polypeptide having retrograded starch degrading activity. See
15 polynucleotide section herein.

In another preferred aspect, the nucleotide sequence comprises or consists of SEQ ID NO: 1. In another preferred aspect, the nucleotide sequence comprises or consists of the mature polypeptide coding sequence of SEQ ID NO: 1. In another preferred aspect, the nucleotide sequence comprises or consists of nucleotides 55 to 1214 of SEQ ID NO: 1.

20 In a further embodiment the nucleotide sequences encode polypeptides comprising or consisting of the amino acid sequence of SEQ ID NO: 2 or the mature polypeptide thereof, which differ from SEQ ID NO: 1 or the mature polypeptide coding sequence thereof by virtue of the degeneracy of the genetic code. The present invention also relates to subsequences of SEQ ID NO: 1 that encode fragments of SEQ ID NO: 2 having retrograded starch degrading
25 activity.

In another aspect, the X143 polypeptide comprises a substitution, deletion, and/or insertion of one or more (or several) amino acids of the mature polypeptide of SEQ ID NO: 2, or a homologous sequence thereof. Preferably, amino acid changes are of a minor nature, that is conservative amino acid substitutions or insertions that do not significantly affect the folding
30 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.

35 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, 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 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: 2152-2156; WO 95/17413; or WO 95/22625. Other methods that can be used include error-prone PCR, phage display (e.g., Lowman *et al.*, 1991, *Biochemistry* 30: 10832-10837; U.S. Patent No. 5,223,409; WO 92/06204), and region-directed mutagenesis (Derbyshire *et al.*, 1986, *Gene* 46: 145; Ner *et al.*, 1988, *DNA* 7: 127).

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

The total number of amino acid substitutions, deletions and/or insertions of the mature polypeptide of SEQ ID NO: 2, is not more than 10, e.g., 1, 2, 3, 4, 5, 6, 7, 8 or 9.

A polypeptide having retrograded starch degrading activity may be obtained from *Aspergillus nidulans*

5 Strains of this 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 (DSM), Centraalbureau Voor Schimmelcultures (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

10 In another more preferred aspect, the polypeptide is an *Aspergillus nidulans* polypeptide having retrograded starch degrading activity. In a most preferred aspect, the polypeptide is an *Aspergillus nidulans* FGSC A1000 polypeptide having retrograded starch degrading activity, e.g., the polypeptide comprising the mature polypeptide of SEQ ID NO: 2.

15 Compositions

In one aspect the invention relates to a composition comprising: (a) an X143 polypeptide having retrograded starch degrading activity and (b) a sulfur-containing compound, and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the combination of the X143 polypeptide having retrograded starch degrading activity and the
20 sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, enhances degradation of retrograded starch material by the X143 polypeptide.

In a particular embodiment the sulfur-containing compound is L-cysteine. In a further particular embodiment the heterocyclic compound (reducing agent) is (1,2-dihydroxyethyl)-3,4-dihydroxyfuran-2(5H)-one (ascorbic acid). In another particular embodiment the sulfur-
25 containing compound is sodium bisulfite (NaHSO₃). In another particular embodiment the cellobiose dehydrogenase is a *Humicola insolens* cellobiose dehydrogenase.

In one embodiment the composition further comprises one or more additional enzymes, e.g., amylolytic enzymes.

30 The one or more additional enzymes may be selected from the group consisting of an acid alpha-amylase, a glucoamylase, a pullulanase, a protease, and a phytase. In one particular embodiment the composition further comprises a glucoamylase. In another embodiment the composition further comprises an acid alpha-amylase and a glucoamylase.

Also contemplated are compositions for starch conversion purposes, which besides the
35 X143 polypeptide and a sulfur containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, may also comprise a glucoamylase, pullulanase, and an acid alpha-amylases.

The composition may also further comprise one or more (*e.g.*, several) proteins selected from the group consisting of a protease, and a phytase. In one aspect, the 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 composition comprises one or more (*e.g.*, several) cellulolytic enzymes. In another aspect, the composition comprises or further comprises one or more (*e.g.*, several) hemicellulolytic enzymes. In another aspect, the composition comprises one or more (*e.g.*, several) cellulolytic enzymes and one or more (*e.g.*, several) hemicellulolytic enzymes. In another aspect, the composition comprises one or more (*e.g.*, several) enzymes selected from the group of cellulolytic enzymes and hemicellulolytic enzymes. In another aspect, the composition comprises an endoglucanase. In another aspect, the composition comprises a cellobiohydrolase. In another aspect, the composition comprises a beta-glucosidase. In another aspect, the composition comprises a polypeptide having cellulolytic enhancing activity. In another aspect, the composition comprises an endoglucanase and a polypeptide having cellulolytic enhancing activity. In another aspect, the composition comprises a cellobiohydrolase and a polypeptide having cellulolytic enhancing activity. In another aspect, the composition comprises a beta-glucosidase and a polypeptide having cellulolytic enhancing activity. In another aspect, the composition comprises an endoglucanase and a cellobiohydrolase. In another aspect, the composition comprises an endoglucanase and a beta-glucosidase. In another aspect, the composition comprises a cellobiohydrolase and a beta-glucosidase. In another aspect, the composition comprises an endoglucanase, a cellobiohydrolase, and a polypeptide having cellulolytic enhancing activity. In another aspect, the composition comprises an endoglucanase, a beta-glucosidase, and a polypeptide having cellulolytic enhancing activity. In another aspect, the composition comprises a cellobiohydrolase, a beta-glucosidase, and a polypeptide having cellulolytic enhancing activity. In another aspect, the composition comprises an endoglucanase, a cellobiohydrolase, and a beta-glucosidase. In another aspect, the composition comprises an endoglucanase, a cellobiohydrolase, a beta-glucosidase, and a polypeptide having cellulolytic enhancing activity.

One or more (*e.g.*, several) components of the 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 composition. One or more (*e.g.*, several) components of the composition may be produced as monocomponents, which are then combined to form the composition. The 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 crude fermentation broth with or without cells removed, a cell lysate with or without cellular debris, a semi-purified or purified enzyme preparation, or a host cell as a source of the enzymes. The enzyme composition may be a dry powder or granulate, a non-dusting granulate, a liquid, a stabilized liquid, or a stabilized protected enzyme. Liquid enzyme preparations may, for instance, be stabilized by adding stabilizers such as a sugar, a sugar alcohol or another polyol, and/or lactic acid or another organic acid according to established processes.

The enzymes can be derived or obtained from any suitable origin, including, bacterial, fungal, yeast, plant, or mammalian origin. The term "obtained" means herein that the enzyme may have been isolated from an organism that naturally produces the enzyme as a native enzyme. 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.

Chemically modified or protein engineered mutants of the 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 amylolytic enzymes may also be prepared by purifying such a protein from a fermentation broth.

In one aspect, the one or more (e.g., several) amylolytic enzymes comprise a commercial amylolytic enzyme preparation.

Amylolytic enzymes useful in the methods and compositions according to the invention.

5 The enzyme(s) and polypeptides described below are to be used in an “effective amount” in processes of the present invention.

Alpha-Amylases

10 Any alpha-amylase may be used, such as of fungal, bacterial or plant origin. In a preferred embodiment the alpha-amylase is an acid alpha-amylase, e.g., acid fungal or acid bacterial alpha-amylase. The term “acid alpha-amylase” means an alpha-amylase (EC 3.2.1.1) which added in an effective amount has activity optimum at a pH in the range of 3 to 7, preferably from 3.5 to 6, or more preferably from 4-5.

15 Bacterial Alpha-Amylases

 An alpha-amylase for use in the present invention may be a bacterial alpha-amylase, e.g., derived from *Bacillus*. In a preferred embodiment the *Bacillus* alpha-amylase is derived from a strain of *Bacillus amyloliquefaciens*, *Bacillus licheniformis*, *Bacillus stearothermophilus*, or *Bacillus subtilis*, but may also be derived from other *Bacillus* sp.

20 Specific examples of alpha-amylases include the *Bacillus amyloliquefaciens* alpha-amylase of SEQ ID NO: 5 in WO 99/19467, the *Bacillus licheniformis* alpha-amylase of SEQ ID NO: 4 in WO 99/19467, and the *Bacillus stearothermophilus* alpha-amylase of SEQ ID NO: 3 in WO 99/19467 (all sequences are hereby incorporated by reference). In an embodiment the alpha-amylase may be an enzyme having a degree of identity of at least 60%, e.g., at least
25 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% to any of the sequences shown in SEQ ID NOS: 3, 4 or 5, respectively, in WO 99/19467.

 The *Bacillus* alpha-amylase may also be a variant and/or hybrid, especially one described in any of WO 96/23873, WO 96/23874, WO 97/41213, WO 99/19467, WO 00/60059,
30 and WO 02/10355 (all documents are hereby incorporated by reference). Specific alpha-amylase variants are disclosed in U.S. Patent Nos. 6,093,562, 6,187,576, and 6,297,038 (hereby incorporated by reference) and include *Bacillus stearothermophilus* alpha-amylase (BSG alpha-amylase) variants having a deletion of one or two amino acids at positions R179 to G182, preferably a double deletion disclosed in WO 96/23873 – see, e.g., page 20, lines 1-10
35 (hereby incorporated by reference), preferably corresponding to delta(181-182) compared to the amino acid sequence of *Bacillus stearothermophilus* alpha-amylase set forth in SEQ ID NO: 3 disclosed in WO 99/19467 or the deletion of amino acids R179 and G180 using SEQ ID

NO: 3 in WO 99/19467 for numbering (which reference is hereby incorporated by reference). Even more preferred are *Bacillus* alpha-amylases, especially *Bacillus stearothermophilus* alpha-amylases, which have a double deletion corresponding to delta(181-182) and further comprise a N193F substitution (also denoted I181* + G182* + N193F) compared to the wild-type BSG alpha-amylase amino acid sequence set forth in SEQ ID NO: 3 disclosed in WO 99/19467.

Bacterial Hybrid Alpha-Amylases

The alpha-amylase may be a hybrid alpha-amylase, e.g., an alpha-amylase comprising 445 C-terminal amino acid residues of the *Bacillus licheniformis* alpha-amylase (shown in SEQ ID NO: 4 of WO 99/19467) and the 37 N-terminal amino acid residues of the alpha-amylase derived from *Bacillus amyloliquefaciens* (shown in SEQ ID NO: 5 of WO 99/19467), with one or more, especially all, of the following substitutions:

G48A+T49I+G107A+H156Y+A181T+N190F+I201F+A209V+Q264S (using the *Bacillus licheniformis* numbering in SEQ ID NO: 4 of WO 99/19467). Also preferred are variants having one or more of the following mutations (or corresponding mutations in other *Bacillus* alpha-amylases): H154Y, A181T, N190F, A209V and Q264S and/or the deletion of two residues between positions 176 and 179, preferably the deletion of E178 and G179 (using SEQ ID NO: 5 of WO 99/19467 for position numbering).

In an embodiment, the bacterial alpha-amylase is dosed in an amount of 0.0005-5 KNU per g DS (dry solids), preferably 0.001-1 KNU per g DS, such as around 0.050 KNU per g DS.

Fungal Alpha-Amylases

Fungal alpha-amylases include alpha-amylases derived from a strain of *Aspergillus*, such as, *Aspergillus kawachii*, *Aspergillus niger* and *Aspergillus oryzae* alpha-amylases.

A preferred acidic fungal alpha-amylase is an alpha-amylase which exhibits a high identity, i.e., at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or even 100% identity to the mature part of the amino acid sequence shown in SEQ ID NO: 10 in WO 96/23874.

Another preferred acid alpha-amylase is derived from a strain of *Aspergillus niger*. In a preferred embodiment the acid fungal alpha-amylase is an *Aspergillus niger* alpha-amylase disclosed as "AMYA_ASPNG" in the Swiss-prot/TrEMBL database under the primary accession no. P56271 and described in WO 89/01969 (Example 3 - incorporated by reference).

Other wild-type alpha-amylases include those derived from a strain of *Meripilus* and *Rhizomucor*, preferably a strain of *Meripilus giganteus* or *Rhizomucor pusillus* (WO 2004/055178 which is incorporated herein by reference).

In a preferred embodiment the alpha-amylase is derived from *Aspergillus kawachii* (Kaneko *et al.*, 1996, *J. Ferment. Bioeng.* 81: 292-298, "Molecular-cloning and determination of the nucleotide-sequence of a gene encoding an acid-stable alpha-amylase from *Aspergillus kawachii*"; and further as EMBL: #AB008370).

5 The fungal alpha-amylase may also be a wild-type enzyme comprising a starch-binding domain (SBD) and an alpha-amylase catalytic domain, or a variant thereof.

Fungal Hybrid Alpha-Amylases

In a preferred embodiment the fungal acid alpha-amylase is a hybrid alpha-amylase. Examples of fungal hybrid alpha-amylases include the ones disclosed in WO 2005/003311, U.S. Application Publication No. 2005/0054071 (Novozymes), and WO 2006/069290 (Novozymes), which are hereby incorporated by reference. A hybrid alpha-amylase may comprise an alpha-amylase catalytic domain (CD) and a carbohydrate-binding domain/module (CBM), such as a starch binding domain (SBD), and optionally a linker.

15 Examples of hybrid alpha-amylases include those disclosed in Tables 1 to 5 of the examples in WO 2006/069290 including the variant with the catalytic domain JA118 and *Athelia rolfsii* SBD (SEQ ID NO: 100 in WO 2006/069290), *Rhizomucor pusillus* alpha-amylase with *Athelia rolfsii* AMG linker and SBD (SEQ ID NO: 101 in WO 2006/069290), *Rhizomucor pusillus* alpha-amylase with *Aspergillus niger* glucoamylase linker and SBD (which is disclosed in Table 20 5 as a combination of amino acid sequences SEQ ID NO: 20, SEQ ID NO: 72 and SEQ ID NO: 96 in U.S. application no. 11/316,535) or as V039 in Table 5 in WO 2006/069290, and *Meripilus giganteus* alpha-amylase with *Athelia rolfsii* glucoamylase linker and SBD (SEQ ID NO: 102 in WO 2006/069290). Other hybrid alpha-amylases are listed in Tables 3, 4, 5, and 6 in Example 4 in U.S. application no. 11/316,535 and WO 2006/069290 (which are hereby 25 incorporated by reference).

Other examples of hybrid alpha-amylases include those disclosed in U.S. Patent Application Publication No. 2005/0054071, including those disclosed in Table 3 on page 15, such as *Aspergillus niger* alpha-amylase with *Aspergillus kawachii* linker and starch binding domain.

30 Other alpha-amylases exhibit a high degree of sequence identity to any of above mentioned alpha-amylases, *i.e.*, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or even 100% identity to the mature enzyme sequences disclosed above.

An acid alpha-amylase may be added in an amount of 0.001 to 10 AFAU/g DS, preferably from 0.01 to 5 AFAU/g DS, especially 0.3 to 2 AFAU/g DS or 0.001 to 1 FAU-F/g DS, preferably 0.01 to 1 FAU-F/g DS.

Commercial Alpha-Amylase Products

Preferred commercial compositions comprising alpha-amylase include MYCOLASE™ (DSM), BAN™, TERMAMYL™ SC, FUNGAMYL™, LIQUOZYME™ X, LIQUOZYME™ SC and SAN™ SUPER, SAN™ EXTRA L (Novozymes A/S) and CLARASE™ L-40,000, DEX-LO™, SPEZYME™ FRED, SPEZYME™ AA, SPEZYME™ ALPHA, SPEZYME™ DELTA AA, GC358, GC980, SPEZYME™ CL and SPEZYME™ RSL (Danisco A/S), and the acid fungal alpha-amylase from *Aspergillus niger* referred to as SP288 (available from Novozymes A/S, Denmark).

Carbohydrate-Source Generating Enzymes (Saccharifying Enzymes)

The term “carbohydrate-source generating enzyme” includes glucoamylase (a glucose generator), beta-amylase and maltogenic amylase (both maltose generators) and also alpha-glucosidase, isoamylase and pullulanase. A carbohydrate-source generating enzyme is capable of producing a carbohydrate that can be used as an energy-source by the fermenting organism(s) in question, for instance, when used in a process of the invention for producing a fermentation product, such as ethanol. The generated carbohydrate may be converted directly or indirectly to the desired fermentation product, preferably ethanol. A mixture of carbohydrate-source generating enzymes may be used. Blends include mixtures comprising at least a glucoamylase and an alpha-amylase, especially an acid amylase, even more preferred an acid fungal alpha-amylase.

The ratio between glucoamylase activity (AGU) and acid fungal alpha-amylase activity (FAU-F) (*i.e.*, AGU per FAU-F) may in a preferred embodiment of the invention be between 0.1 and 100 AGU/FAU-F, in particular between 2 and 50 AGU/FAU-F, such as in the range from 10-40 AGU/FAU-F, especially when performing a one-step fermentation (raw starch hydrolysis - RSH), *i.e.*, when saccharification and fermentation are carried out simultaneously (*i.e.*, without a liquefaction step).

In a conventional starch-to-ethanol process (*i.e.*, including a liquefaction step) the ratio may preferably be as defined in EP 140410, especially when saccharification and fermentation are carried out simultaneously.

Glucoamylases

The term “glucoamylase” (1,4-alpha-D-glucan glucohydrolase, EC 3.2.1.3) is an enzyme, which catalyzes the release of D-glucose from the non-reducing ends of starch or related oligo- and polysaccharide molecules.

The glucoamylase may added in an amount of 0.001 to 10 AGU/g DS, preferably from 0.01 to 5 AGU/g DS, such as around 0.1, 0.3, 0.5, 1 or 2 AGU/g DS, especially 0.1 to 0.5 AGU/g DS or 0.02-20 AGU/g DS, preferably 0.1-10 AGU/g DS.

A glucoamylase may be derived from any suitable source, e.g., derived from a microorganism or a plant. Preferred glucoamylases are of fungal or bacterial origin, selected from the group consisting of *Aspergillus* glucoamylases, in particular *Aspergillus niger* G1 or G2 glucoamylase (Boel *et al.*, 1984, *EMBO J.* 3(5): 1097-1102), or variants thereof, such as those disclosed in WO 92/00381, WO 00/04136 and WO 01/04273 (from Novozymes, Denmark); the *A. awamori* glucoamylase disclosed in WO 84/02921, *Aspergillus oryzae* glucoamylase (Hata *et al.*, 1991, *Agric. Biol. Chem.* 55(4): 941-949), or variants or fragments thereof. Other *Aspergillus* glucoamylase variants include variants with enhanced thermal stability: G137A and G139A (Chen *et al.*, 1996, *Prot. Eng.* 9: 499-505); D257E and D293E/Q (Chen *et al.*, 1995, *Prot. Eng.* 8: 575-582); N182 (Chen *et al.*, 1994, *Biochem. J.* 301: 275-281); disulphide bonds, A246C (Fierobe *et al.*, 1996, *Biochemistry* 35: 8698-8704; and introduction of Pro residues in positions A435 and S436 (Li *et al.*, 1997, *Prot. Eng.* 10: 1199-1204).

Other glucoamylases include *Athelia rolfsii* (previously denoted *Corticium rolfsii*) glucoamylase (see U.S. Patent No. 4,727,026 and Nagasaka *et al.*, 1998, *Appl. Microbiol. Biotechnol.* 50: 323-330), *Talaromyces* glucoamylases, in particular derived from *Talaromyces duponti*, *Talaromyces emersonii* (WO 99/28448), *Talaromyces leycettanus* (U.S. Patent No. Re. 32,153), and *Talaromyces thermophilus* (U.S. Patent No. 4,587,215).

Bacterial glucoamylases include glucoamylases from *Clostridium*, in particular *C. thermoamylolyticum* (EP 135138) and *C. thermohydrosulfuricum* (WO 86/01831), *Trametes cingulata*, *Pachykytospora papyracea*, and *Leucopaxillus giganteus*, all disclosed in WO 2006/069289; or *Peniophora rufomarginata* disclosed in PCT/US2007/066618; or a mixture thereof. A hybrid glucoamylase may be used in the present invention. Examples of hybrid glucoamylases are disclosed in WO 2005/045018. Specific examples include the hybrid glucoamylase disclosed in Tables 1 and 4 of Example 1 (which hybrids are hereby incorporated by reference).

The glucoamylase may have a high degree of sequence identity to any of above mentioned glucoamylases, *i.e.*, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or even 100% identity to the mature enzymes sequences mentioned above.

Commercially available glucoamylase compositions include AMG 200L; AMG 300L; SAN™ SUPER, SAN™ EXTRA L, SPIRIZYME™ PLUS, SPIRIZYME™ FUEL, SPIRIZYME™ B4U, SPIRIZYME ULTRA™ and AMG™ E (from Novozymes A/S, Denmark); OPTIDEX™ 300, GC480™ and GC147™ (from Genencor Int., USA); AMIGASE™ and AMIGASE™ PLUS (from DSM); G-ZYME™ G900, G-ZYME™ and G990 ZR (from Genencor Int.). [

Glucoamylases may be added in an amount of 0.02-20 AGU/g DS, preferably 0.1-10 AGU/g DS, especially between 1-5 AGU/g DS, such as 0.1-2 AGU/g DS, such as 0.5 AGU/g

DS or in an amount of 0.0001-20 AGU/g DS, preferably 0.001-10 AGU/g DS, especially between 0.01-5 AGU/g DS, such as 0.1-2 AGU/g DS.

Beta-amylases

5 A beta-amylase (E.C 3.2.1.2) is the name traditionally given to exo-acting maltogenic amylases, which catalyze the hydrolysis of 1,4-alpha-glucosidic linkages in amylose, amylopectin and related glucose polymers. Maltose units are successively removed from the non-reducing chain ends in a step-wise manner until the molecule is degraded or, in the case of amylopectin, until a branch point is reached. The maltose released has the beta anomeric
10 configuration, hence the name beta-amylase.

Beta-amylases have been isolated from various plants and microorganisms (Fogarty and Kelly, 1979, *Progress in Industrial Microbiology* 15: 112-115). These beta-amylases are characterized by having a temperature optimum in the range from 40°C to 65°C and a pH optimum in the range from 4.5 to 7. A commercially available beta-amylase from barley is
15 NOVOZYM™ WBA from Novozymes A/S, Denmark and SPEZYME™ BBA 1500 from Genencor Int., USA.

Maltogenic Amylases

The amylase may also be a maltogenic alpha-amylase (glucan 1,4-alpha-maltohydrolase, EC 3.2.1.133), which catalyzes the hydrolysis of amylose and amylopectin to
20 maltose in the alpha-configuration. A maltogenic amylase from *Bacillus stearothermophilus* strain NCIB 11837 is commercially available from Novozymes A/S. Maltogenic alpha-amylases are described in U.S. Patent Nos. 4,598,048, 4,604,355 and 6,162,628, which are hereby incorporated by reference.

25 The maltogenic amylase may be added in an amount of 0.05-5 mg total protein/gram DS or 0.05-5 MANU/g DS.

Phytases

Any phytase may be used in a process of the present invention. Phytases are enzymes
30 that degrade phytates and/or phytic acid by specifically hydrolyzing the ester link between inositol and phosphorus. Phytase activity is credited with phosphorus and ion availability in many ingredients. In some embodiments, the phytase is capable of liberating at least one inorganic phosphate from an inositol hexaphosphate (e.g., phytic acid). Phytases can be grouped according to their preference for a specific position of the phosphate ester group on
35 the phytate molecule at which hydrolysis is initiated (e.g., 3-phytase (EC 3.1.3.8) or 6-phytase (EC 3.1.3.26)). An example of phytase is myo-inositol-hexakiphosphate-3-phosphohydrolase.

Phytases can be obtained from microorganisms such as fungal and bacterial organisms. For example, the phytase may be obtained from filamentous fungi such as *Aspergillus* (e.g., *A. ficuum*, *A. fumigatus*, *A. niger*, and *A. terreus*), *Cladospirum*, *Mucor* (e.g., *Mucor piriformis*), *Myceliophthora* (e.g., *M. thermophila*), *Penicillium* (e.g., *P. hordei* (ATCC No. 22053)), *P. piceum* (ATCC No. 10519), or *P. brevi-compactum* (ATCC No. 48944), *Talaromyces* (e.g., *T. thermophilus*), *Thermomyces* (WO 99/49740), and *Trichoderma* spp. (e.g., *T. reesei*).

In an embodiment, the phytate-degrading enzyme is obtained from yeast (e.g., *Arxula adenivorans*, *Pichia anomala*, *Schwanniomyces occidentalis*), gram-negative bacteria (e.g., *Escherichia coli*, *Klebsiella* spp., *Pseudomonas* spp.), and gram-positive bacteria (e.g., *Bacillus* spp. such as *Bacillus subtilis*).

The phytase also may be obtained from *Citrobacter*, *Enterbacter*, or *Peniophora*.

In an embodiment, the phytase is derived from *Buttiauxiella* spp. such as *B. agrestis*, *B. brennerae*, *B. ferragutiase*, *B. gaviniae*, *B. izardii*, *B. noackiae*, and *B. warmboldiae*. In some embodiments, the phytase is a phytase disclosed in WO 2006/043178 or U.S. application no. 11/714,487.

In one preferred embodiment, the phytase has at least 75%, at least 80%, at least 85%, at least 88%, at least 90%, at least 93%, at least 95%, at least 96%, at least 97%, at least 98% and at least 99% sequence identity to the amino acid sequence set forth in SEQ ID NO: 31 of U.S. Application No. 12/263,886.

Commercially-available phytases are NATUPHOS (BASF), RONOZYME P (Novozymes A/S), PHZYME (Danisco A/S, Diversa) and FINASE (AB Enzymes). The method for determining microbial phytase activity and the definition of a phytase unit is disclosed in Engelen *et al.*, 1994, *Journal of AOAC International* 77: 760-764. The phytase may be a wild-type phytase, an active variant or active fragment thereof.

Pullulanases

Any pullulanase may be used in a process of the present invention. In an embodiment, the pullulanase is a GH57 pullulanase, e.g., a pullulanase obtained from a strain of *Thermococcus*, including *Thermococcus* sp. AM4, *Thermococcus* sp. HJ21, *Thermococcus barophilus*, *Thermococcus gammatolerans*, *Thermococcus hydrothermalis*, *Thermococcus kodakarensis*, *Thermococcus litoralis*, and *Thermococcus onnurineus*; or from a strain of *Pyrococcus*, such as *Pyrococcus abyssi* and *Pyrococcus furiosus*.

Proteases

A protease may be added during saccharification, fermentation, simultaneous saccharification and fermentation. The protease may be any protease. In a preferred embodiment

the protease is an acid protease of microbial origin, preferably of fungal or bacterial origin. An acid fungal protease is preferred, but also other proteases can be used.

Suitable proteases include microbial proteases, such as fungal and bacterial proteases. Preferred proteases are acidic proteases, *i.e.*, proteases characterized by the ability to hydrolyze proteins under acidic conditions below pH 7.

The acid fungal protease may be derived from *Aspergillus*, *Candida*, *Coriolus*, *Endothia*, *Entomophtra*, *Irpex*, *Mucor*, *Penicillium*, *Rhizopus*, *Sclerotium*, and *Torulopsis*. In particular, the protease may be derived from *Aspergillus aculeatus* (WO 95/02044), *Aspergillus awamori* (Hayashida *et al.*, 1977, *Agric. Biol. Chem.* 42(5), 927-933), *Aspergillus niger* (see, *e.g.*, Koaze *et al.*, 1964, *Agr. Biol. Chem. Japan* 28: 216), *Aspergillus saitoi* (see, *e.g.*, Yoshida, 1954, *J. Agr. Chem. Soc. Japan* 28: 66), or *Aspergillus oryzae*, such as the pepA protease; and acidic proteases from *Mucor miehei* or *Mucor pusillus*.

The protease may be a neutral or alkaline protease, such as a protease derived from a strain of *Bacillus*. A particular protease is derived from *Bacillus amyloliquefaciens* and has the sequence obtainable at the Swissprot Database, Accession no. P06832. The proteases may have at least 90% sequence identity to the amino acid sequence disclosed in the Swissprot Database, Accession no. P06832 such as at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, or particularly at least 99% identity.

The protease may have at least 90% sequence identity to the amino acid sequence disclosed as SEQ ID NO: 1 in WO 2003/048353 such as at 92%, at least 95%, at least 96%, at least 97%, at least 98%, or particularly at least 99% identity.

The protease may be a papain-like protease selected from the group consisting of proteases within EC 3.4.22.* (cysteine protease), such as EC 3.4.22.2 (papain), EC 3.4.22.6 (chymopapain), EC 3.4.22.7 (asclepain), EC 3.4.22.14 (actinidain), EC 3.4.22.15 (cathepsin L), EC 3.4.22.25 (glycyl endopeptidase) and EC 3.4.22.30 (caricain).

In an embodiment, the protease is a protease preparation derived from a strain of *Aspergillus*, such as *Aspergillus oryzae*. In another embodiment the protease is derived from a strain of *Rhizomucor*, preferably *Rhizomucor miehei*. In another embodiment the protease is a protease preparation, preferably a mixture of a proteolytic preparation derived from a strain of *Aspergillus*, such as *Aspergillus oryzae*, and a protease derived from a strain of *Rhizomucor*, preferably *Rhizomucor miehei*.

Aspartic acid proteases are described in, for example, Handbook of Proteolytic Enzymes, Edited by A.J. Barrett, N.D. Rawlings and J.F. Woessner, Academic Press, San Diego, 1998, Chapter 270. Examples of aspartic acid proteases include, *e.g.*, those disclosed in Berka *et al.*, 1990, *Gene* 96: 313; Berka *et al.*, 1993, *Gene* 125: 195-198; and Gomi *et al.*, 1993, *Biosci. Biotech. Biochem.* 57: 1095-1100, which are hereby incorporated by reference.

The protease also may be a metalloprotease, which is defined as a protease selected from the group consisting of:

(a) proteases belonging to EC 3.4.24 (metalloendopeptidases); preferably EC 3.4.24.39 (acid metallo proteinases);

(b) metalloproteases belonging to the M group of the above Handbook;

(c) metalloproteases not yet assigned to clans (designation: Clan MX), or belonging to either one of clans MA, MB, MC, MD, ME, MF, MG, MH (as defined at pp. 989-991 of the above Handbook);

(d) other families of metalloproteases (as defined at pp. 1448-1452 of the above Handbook);

(e) metalloproteases with a HEXXH motif;

(f) metalloproteases with an HEFTH motif;

(g) metalloproteases belonging to either one of families M3, M26, M27, M32, M34, M35, M36, M41, M43, or M47 (as defined at pp. 1448-1452 of the above Handbook);

(h) metalloproteases belonging to the M28E family; and

(i) metalloproteases belonging to family M35 (as defined at pp. 1492-1495 of the above Handbook).

In other particular embodiments, metalloproteases are hydrolases in which the nucleophilic attack on a peptide bond is mediated by a water molecule, which is activated by a divalent metal cation. Examples of divalent cations are zinc, cobalt or manganese. The metal ion may be held in place by amino acid ligands. The number of ligands may be five, four, three, two, one or zero. In a particular embodiment the number is two or three, preferably three.

There are no limitations on the origin of the metalloprotease used in a process of the invention. In an embodiment the metalloprotease is classified as EC 3.4.24, preferably EC 3.4.24.39. In one embodiment, the metalloprotease is an acid-stable metalloprotease, *e.g.*, a fungal acid-stable metalloprotease, such as a metalloprotease derived from a strain of the genus *Thermoascus*, preferably a strain of *Thermoascus aurantiacus*, especially *Thermoascus aurantiacus* CGMCC No. 0670 (classified as EC 3.4.24.39). In another embodiment, the metalloprotease is derived from a strain of the genus *Aspergillus*, preferably a strain of *Aspergillus oryzae*.

In one embodiment the metalloprotease has a degree of sequence identity to amino acids -178 to 177, -159 to 177, or preferably amino acids 1 to 177 (the mature polypeptide) of SEQ ID NO: 1 of WO 2010/008841 (a *Thermoascus aurantiacus* metalloprotease) of at least 80%, at least 82%, at least 85%, at least 90%, at least 95%, or at least 97%; and which have metalloprotease activity. In particular embodiments, the metalloprotease consists of an amino acid sequence with a degree of identity to SEQ ID NO: 1 as mentioned above.

The *Thermoascus aurantiacus* metalloprotease is a preferred example of a metalloprotease suitable for use in a process of the invention. Another metalloprotease is derived from *Aspergillus oryzae* and comprises the sequence of SEQ ID NO: 11 disclosed in WO 2003/048353, or amino acids -23-353; -23-374; -23-397; 1-353; 1-374; 1-397; 177-353; 177-374; or 177-397 thereof, and SEQ ID NO: 10 disclosed in WO 2003/048353.

Another metalloprotease suitable for use in a process of the invention is the *Aspergillus oryzae* metalloprotease comprising SEQ ID NO: 5 of WO 2010/008841, or a metalloprotease is an isolated polypeptide which has a degree of identity to SEQ ID NO: 5 of at least about 80%, at least 82%, at least 85%, at least 90%, at least 95%, or at least 97%; and which have metalloprotease activity. In particular embodiments, the metalloprotease consists of the amino acid sequence of SEQ ID NO: 5.

In a particular embodiment, a metalloprotease has an amino acid sequence that differs by forty, thirty-five, thirty, twenty-five, twenty, or by fifteen amino acids from amino acids -178 to 177, -159 to 177, or +1 to 177 of the amino acid sequences of the *Thermoascus aurantiacus* or *Aspergillus oryzae* metalloprotease.

In another embodiment, a metalloprotease has an amino acid sequence that differs by ten, or by nine, or by eight, or by seven, or by six, or by five amino acids from amino acids -178 to 177, -159 to 177, or +1 to 177 of the amino acid sequences of these metalloproteases, e.g., by four, by three, by two, or by one amino acid.

In particular embodiments, the metalloprotease a) comprises or b) consists of

- i) the amino acid sequence of amino acids -178 to 177, -159 to 177, or +1 to 177 of SEQ ID NO:1 of WO 2010/008841;
- ii) the amino acid sequence of amino acids -23-353, -23-374, -23-397, 1-353, 1-374, 1-397, 177-353, 177-374, or 177-397 of SEQ ID NO: 3 of WO 2010/008841;
- iii) the amino acid sequence of SEQ ID NO: 5 of WO 2010/008841; or allelic variants, or fragments, of the sequences of i), ii), and iii) that have protease activity.

A fragment of amino acids -178 to 177, -159 to 177, or +1 to 177 of SEQ ID NO: 1 of WO 2010/008841 or of amino acids -23-353, -23-374, -23-397, 1-353, 1-374, 1-397, 177-353, 177-374, or 177-397 of SEQ ID NO: 3 of WO 2010/008841; is a polypeptide having one or more amino acids deleted from the amino and/or carboxyl terminus of these amino acid sequences. In one embodiment a fragment contains at least 75 amino acid residues, or at least 100 amino acid residues, or at least 125 amino acid residues, or at least 150 amino acid residues, or at least 160 amino acid residues, or at least 165 amino acid residues, or at least 170 amino acid residues, or at least 175 amino acid residues.

In another embodiment, the metalloprotease is combined with another protease, such as a fungal protease, preferably an acid fungal protease.

Commercially available products include ALCALASE®, ESPERASE™, FLAVOURZYME™, NEUTRASE®, RENNILASE®, NOVOZYM™ FM 2.0L, and iZyme BA (available from Novozymes A/S, Denmark) and GC106™ and SPEZYME™ FAN from Genencor International, Inc., USA.

The protease may be present in an amount of 0.0001-1 mg enzyme protein per g DS, preferably 0.001 to 0.1 mg enzyme protein per g DS. Alternatively, the protease may be present in an amount of 0.0001 to 1 LAPU/g DS, preferably 0.001 to 0.1 LAPU/g DS and/or 0.0001 to 1 mAU-RH/g DS, preferably 0.001 to 0.1 mAU-RH/g DS.

Cellulolytic enzymes useful in the compositions and methods according to the invention

In one aspect, the one or more (e.g., several) cellulolytic enzymes comprise a commercial cellulolytic enzyme preparation. Examples of commercial cellulolytic enzyme preparations suitable for use in the present invention include, for example, CELLIC® CTec (Novozymes A/S), CELLIC® CTec2 (Novozymes A/S), CELLUCLAST™ (Novozymes A/S), 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 (Röhm GmbH), FIBREZYME® LDI (Dyadic International, Inc.), FIBREZYME® LBR (Dyadic International, Inc.), or VISCOSTAR® 150L (Dyadic International, Inc.). The cellulase enzymes are added in amounts effective from about 0.001 to about 5.0 wt % of solids, e.g., about 0.025 to about 4.0 wt % of solids or about 0.005 to about 2.0 wt % of solids.

Examples of bacterial endoglucanases that can be used in the processes 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 2005/093050); *Thermobifida fusca* endoglucanase III (WO 2005/093050); and *Thermobifida fusca* endoglucanase V (WO 2005/093050).

Examples of fungal endoglucanases that can be used in the present invention, include, but are not limited to, a *Trichoderma reesei* endoglucanase I (Penttilä *et al.*, 1986, *Gene* 45: 253-263, *Trichoderma reesei* Cel7B endoglucanase I (GENBANK™ accession no. M15665), *Trichoderma reesei* endoglucanase II (Saloheimo, *et al.*, 1988, *Gene* 63:11-22), *Trichoderma reesei* Cel5A endoglucanase II (GENBANK™ accession no. M19373), *Trichoderma reesei* endoglucanase III (Okada *et al.*, 1988, *Appl. Environ. Microbiol.* 64: 555-563, GENBANK™ accession no. AB003694), *Trichoderma reesei* endoglucanase V (Saloheimo *et al.*, 1994, *Molecular Microbiology* 13: 219-228, GENBANK™ accession no. Z33381), *Aspergillus aculeatus* endoglucanase (Ooi *et al.*, 1990, *Nucleic Acids Research* 18: 5884), *Aspergillus kawachii* endoglucanase (Sakamoto *et al.*, 1995, *Current Genetics* 27: 435-439), *Erwinia*

carotovora endoglucanase (Saarilahti *et al.*, 1990, *Gene* 90: 9-14), *Fusarium oxysporum* endoglucanase (GENBANK™ accession no. L29381), *Hemicola grisea* var. *thermoidea* endoglucanase (GENBANK™ accession no. AB003107), *Melanocarpus albomyces* endoglucanase (GENBANK™ accession no. MAL515703), *Neurospora crassa* endoglucanase (GENBANK™ accession no. XM_324477), *Hemicola insolens* endoglucanase V, *Myceliophthora thermophila* CBS 117.65 endoglucanase, basidiomycete CBS 495.95 endoglucanase, basidiomycete CBS 494.95 endoglucanase, *Thielavia terrestris* NRRL 8126 CEL6B endoglucanase, *Thielavia terrestris* NRRL 8126 CEL6C endoglucanase, *Thielavia terrestris* NRRL 8126 CEL7C endoglucanase, *Thielavia terrestris* NRRL 8126 CEL7E endoglucanase, *Thielavia terrestris* NRRL 8126 CEL7F endoglucanase, *Cladorrhinum foecundissimum* ATCC 62373 CEL7A endoglucanase, and *Trichoderma reesei* strain No. VTT-D-80133 endoglucanase (GENBANK™ accession no. M15665).

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

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

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

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

Other cellulolytic enzymes that may be used in the present invention are described in WO 98/13465, WO 98/015619, WO 98/015633, WO 99/06574, WO 99/10481, WO 99/025847, WO 99/031255, WO 02/101078, WO 03/027306, WO 03/052054, WO 03/052055, WO

03/052056, WO 03/052057, WO 03/052118, WO 2004/016760, WO 2004/043980, WO 2004/048592, WO 2005/001065, WO 2005/028636, WO 2005/093050, WO 2005/093073, WO 2006/074005, WO 2006/117432, WO 2007/071818, WO 2007/071820, WO 2008/008070, WO 2008/008793, U.S. Patent No. 5,457,046, U.S. Patent No. 5,648,263, and U.S. Patent No. 5,686,593.

In the processes of the present invention, any GH61 polypeptide having cellulolytic enhancing activity can be used as a component of the enzyme composition.

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

Nucleic Acid Constructs

An isolated polynucleotide encoding a polypeptide, e.g., a polypeptide having retrograded starch degrading activity, may be manipulated in a variety of ways to provide for expression of the polypeptide by constructing a nucleic acid construct comprising an isolated polynucleotide encoding the polypeptide 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. Manipulation of the polynucleotide's sequence prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotide sequences utilizing recombinant DNA methods are well known in the art.

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

Examples of suitable promoters for directing the transcription of the nucleic acid constructs in the present invention in a bacterial host cell are the promoters obtained from the *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus stearothermophilus*

maltogenic amylase gene (*amyM*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus subtilis* *xylA* and *xylB* genes, *E. coli* *lac* operon, *Streptomyces coelicolor* agarase gene (*dagA*), and prokaryotic beta-lactamase gene (Villa-Kamaroff *et al.*, 1978, *Proc. Natl. Acad. Sci. USA* 75: 3727-3731), as well as the *tac* promoter (DeBoer *et al.*, 1983, *Proc. Natl. Acad. Sci. USA* 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in Gilbert *et al.*, 1980, *Scientific American*, 242: 74-94; and in Sambrook *et al.*, 1989, *supra*.

Examples of suitable promoters for directing the transcription of the nucleic acid constructs in the present invention in a filamentous fungal host cell are promoters obtained from the genes for *Aspergillus nidulans* acetamidase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Aspergillus oryzae* TKA amylase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Fusarium oxysporum* trypsin-like protease (WO 96/00787), *Fusarium venenatum* amyloglucosidase (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), *Rhizomucor miehei* lipase, *Rhizomucor miehei* aspartic proteinase, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei* xylanase II, *Trichoderma reesei* beta-xylosidase, as well as the NA2-tpi promoter (a modified promoter from a gene encoding a neutral alpha-amylase in *Aspergilli* in which the untranslated leader has been replaced by an untranslated leader from a gene encoding triose phosphate isomerase in *Aspergilli*; non-limiting examples include modified promoters from the gene encoding neutral alpha-amylase in *Aspergillus niger* in which the untranslated leader has been replaced by an untranslated leader from the gene encoding triose phosphate isomerase in *Aspergillus nidulans* or *Aspergillus oryzae*); and mutant, truncated, and hybrid promoters thereof.

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

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

Preferred terminators for 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.

5 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*.

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

15 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).

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

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

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

30 The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a polypeptide and directs the polypeptide into the cell's secretory pathway. The 5'-end of the coding sequence of the polynucleotide may inherently contain a signal peptide coding sequence naturally linked in translation reading frame with the segment of the coding sequence that encodes the polypeptide. Alternatively, the 5'-end of the coding sequence may contain a signal peptide coding sequence that is foreign to the coding
35 sequence. The foreign signal peptide coding sequence may be required where the coding sequence does not naturally contain a signal peptide coding sequence. Alternatively, the foreign signal peptide coding sequence may simply replace the natural signal peptide coding

sequence in order to enhance secretion of the polypeptide. However, any signal peptide coding sequence that directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used.

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

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

It may also be desirable to add regulatory sequences that allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those that cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems include the *lac*, *tac*, and *trp* operator systems. In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the *Aspergillus niger* glucoamylase promoter, *Aspergillus oryzae* TAKA alpha-amylase promoter, and *Aspergillus oryzae* glucoamylase promoter may be used. Other examples of regulatory sequences are those that allow for gene amplification. In eukaryotic systems, these regulatory sequences

include the dihydrofolate reductase gene that is amplified in the presence of methotrexate, and the metallothionein genes that are amplified with heavy metals. In these cases, the polynucleotide encoding the polypeptide would be operably linked with the regulatory sequence.

5

Expression Vectors

The various nucleotide and control sequences may be joined together to produce a recombinant expression vector that may include one or more (e.g., several) convenient restriction sites to allow for insertion or substitution of a polynucleotide encoding a polypeptide, e.g., a polypeptide having cellulolytic enhancing activity, a cellulolytic enzyme, a hemicellulolytic enzyme, etc., at such sites. Alternatively, the polynucleotide may be expressed by inserting the polynucleotide or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

15

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.

20

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.

25

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.

30

Examples of bacterial selectable markers are the *dal* genes from *Bacillus subtilis* or *Bacillus licheniformis*, or markers that confer antibiotic resistance such as ampicillin, chloramphenicol, kanamycin, or tetracycline resistance. Suitable markers for yeast host cells are ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. Selectable markers for use in a filamentous fungal host cell include, but are not limited to, *amdS* (acetamidase), *argB* (ornithine

35

carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hph* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenylyltransferase), and *trpC* (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are the *amdS* and *pyrG* genes of *Aspergillus nidulans* or
5 *Aspergillus oryzae* and the *bar* gene of *Streptomyces hygroscopicus*.

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
10 sequence encoding the polypeptide or any other element of the vector for integration into the genome by homologous or non-homologous recombination. Alternatively, the vector may contain additional polynucleotides for directing integration by homologous recombination into the genome of the host cell at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should contain a
15 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
20 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
25 cell. The term "origin of replication" or "plasmid replicator" means a polynucleotide that enables a plasmid or vector to replicate *in vivo*.

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

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

Examples of origins of replication useful in a filamentous fungal cell are AMA1 and ANS1 (Gems *et al.*, 1991, *Gene* 98: 61-67; Cullen *et al.*, 1987, *Nucleic Acids Res.* 15: 9163-
35 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 polypeptide. An increase in the copy number of the polynucleotide can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the polynucleotide where cells
5 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*).
10

Methods for Processing Retrograded Starch Material

The composition according to the invention may be used for starch processes, in particular starch conversion, especially saccharification of starch and simultaneous
15 saccharification and fermentation (SSF).

Further, the compositions according to the invention are particularly useful in the production of sweeteners and ethanol (see, e.g., U.S. Patent No. 5,231,017, which is hereby incorporated by reference), such as fuel, drinking and industrial ethanol, from starch or whole grains.

The compositions and methods of the present invention can be used to degrade
20 retrograded starch and improve yield during saccharification of a starch substrate comprising retrograded starch to fermentable sugars and convert the fermentable sugars to many useful substances, e.g., fuel, potable ethanol, and/or fermentation products (e.g., acids, alcohols, ketones, gases, and the like).

The present invention further relates to methods for degrading or converting a
25 retrograded starch material, comprising: treating the retrograded starch material with a composition in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the X143 polypeptide is selected from the group
30 consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing
35 agent), and/or a cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition.

In a particular embodiment the sulfur-containing compound is selected from L-cysteine. In a further particular embodiment the heterocyclic compound (reducing agent) is selected from is (1,2-dihydroxyethyl)-3,4-dihydroxyfuran-2(5H)-one (ascorbic acid). In another particular embodiment the sulfur-containing compound is a reducing agent and is selected from sulfur(IV) oxyanion, particularly sodium bisulfite (NaHSO_3). In another particular embodiment the cellobiose dehydrogenase is a *Humicola insolens* cellobiose dehydrogenase.

In one aspect, the method above further comprises recovering the degraded or converted retrograded starch material. Soluble products of degradation or conversion of the retrograded starch material can be separated from the insoluble cellulosic material using technology well known in the art such as, for example, centrifugation, filtration, and gravity settling.

Starch Processing

Native starch consists of microscopic granules, which are insoluble in water at room temperature. When an aqueous starch slurry is heated, the granules swell and eventually burst, dispersing the starch molecules into the solution. At temperatures up to about 50°C to 75°C the swelling may be reversible. However, with higher temperatures an irreversible swelling called "gelatinization" begins. During this "gelatinization" process there is a dramatic increase in viscosity. Granular starch to be processed may be a highly refined starch quality, preferably at least 90%, at least 95%, at least 97% or at least 99.5% pure or it may be a more crude starch-containing materials comprising (e.g., milled) whole grains including non-starch fractions such as germ residues and fibers. The raw material, such as whole grains, may be reduced in particle size, e.g., by milling, in order to open up the structure and allowing for further processing. In dry milling whole kernels are milled and used. Wet milling gives a good separation of germ and meal (starch granules and protein) and is often applied at locations where the starch hydrolyzate is used in the production of, e.g., syrups. Both dry and wet milling are well known in the art of starch processing and may be used in a process of the invention. Methods for reducing the particle size of the starch containing material are well known to those skilled in the art.

As the solids level is 30-40% in a typical industrial process, the starch has to be thinned or "liquefied" so that it can be suitably processed. This reduction in viscosity is primarily attained by enzymatic degradation in current commercial practice.

Liquefaction is carried out in the presence of an alpha-amylase, preferably a bacterial alpha-amylase and/or acid fungal alpha-amylase. In an embodiment, a phytase and/or a protease is also present during liquefaction. In an embodiment, viscosity reducing enzymes such as a xylanase and/or beta-glucanase is also present during liquefaction.

During liquefaction, the long-chained starch is degraded into branched and linear shorter units (maltodextrins) by an alpha-amylase. Liquefaction may be carried out as a three-step hot slurry process. The slurry is heated to between 60-95°C (*e.g.*, 70-90°C, such as 77-86°C, 80-85°C, 83-85°C) and an alpha-amylase is added to initiate liquefaction (thinning).

5 The slurry may in an embodiment be jet-cooked at between 95-140°C, *e.g.*, 105-125°C, for about 1-15 minutes, *e.g.*, about 3-10 minutes, especially around 5 minutes. The slurry is then cooled to 60-95°C and more alpha-amylase is added to obtain final hydrolysis (secondary liquefaction). The jet-cooking process is carried out at pH 4.5-6.5, typically at a pH between 5 and 6. The alpha-amylase may be added as a single dose, *e.g.*, before jet cooking.

10 The liquefaction process is carried out at between 70-95°C, such as 80-90°C, such as around 85°C, for about 10 minutes to 5 hours, typically for 1-2 hours. The pH is between 4 and 7, such as between 5.5 and 6.2. In order to ensure optimal enzyme stability under these conditions, calcium may optionally be added (to provide 1-60 ppm free calcium ions, such as about 40 ppm free calcium ions). After such treatment, the liquefied starch will typically have a
15 "dextrose equivalent" (DE) of 10-15.

Generally liquefaction and liquefaction conditions are well known in the art.

Examples of alpha-amylase are disclosed in the "Alpha-Amylases" section below.

Saccharification may be carried out using conditions well known in the art with a carbohydrate-source generating enzyme, in particular a glucoamylase, or a beta-amylase and
20 optionally a debranching enzyme, such as an isoamylase or a pullulanase. For instance, a full saccharification step may last from about 24 to about 72 hours. However, it is common to do a pre-saccharification of typically 40-90 minutes at a temperature between 30-65°C, typically about 60°C, followed by complete saccharification during fermentation in a simultaneous saccharification and fermentation (SSF) process. Saccharification is typically carried out at a
25 temperature in the range of 20-75°C, *e.g.*, 25-65°C and 40-70°C, typically around 60°C, and at a pH between about 4 and 5, normally at about pH 4.5.

The saccharification and fermentation steps may be carried out either sequentially or simultaneously. In an embodiment, saccharification and fermentation are performed simultaneously (referred to as "SSF"). However, it is common to perform a pre-saccharification
30 step for about 30 minutes to 2 hours (*e.g.*, 30 to 90 minutes) at a temperature of 30 to 65°C, typically around 60°C which is followed by a complete saccharification during fermentation referred to as simultaneous saccharification and fermentation (SSF). The pH is usually between 3-6, more particularly 4-5, most particularly 4.2-4.8, *e.g.*, pH 4.5. In a simultaneous saccharification and fermentation (SSF) process, there is no holding stage for saccharification,
35 rather, the yeast and enzymes are added together. The most widely used process to produce a fermentation product, especially ethanol, is the simultaneous saccharification and fermentation (SSF) process, in which there is no holding stage for the saccharification, meaning that a

fermenting organism, such as yeast, and enzyme(s), may be added together. SSF may typically be carried out at a temperature from 25°C to 40°C, such as from 28°C to 35°C, such as from 30°C to 34°C, preferably around about 32°C. In an embodiment fermentation is ongoing for 6 to 120 hours, in particular 24 to 96 hours.

5 In a typical saccharification process, maltodextrins produced during liquefaction are converted into dextrose by adding a glucoamylase and a debranching enzyme, such as an isoamylase (U.S. Patent No. 4,335,208) or a pullulanase. The temperature is lowered to 60°C, prior to the addition of the glucoamylase and debranching enzyme. The saccharification process proceeds for 24-72 hours. Prior to addition of the saccharifying enzymes, the pH is
10 reduced to below 4.5, while maintaining a high temperature (above 95°C), to inactivate the liquefying alpha-amylase. This process reduces the formation of short oligosaccharide called "panose precursors," which cannot be hydrolyzed properly by the debranching enzyme. Normally, about 0.2-0.5% of the saccharification product is the branched trisaccharide panose (Glc α 1-6Glc α 1-4Glc), which cannot be degraded by a pullulanase. If active amylase from
15 the liquefaction remains present during saccharification (*i.e.*, no denaturing), the amount of panose can be as high as 1-2%, which is highly undesirable since it lowers the saccharification yield significantly.

Other fermentation products may be fermented at conditions and temperatures well known to persons skilled in the art, suitable for the fermenting organism in question.

20 The fermentation product may be recovered by methods well known in the art, *e.g.*, by distillation. Examples of carbohydrate-source generating enzymes, including glucoamylases, are disclosed in the "Enzymes" section below.

In a particular embodiment, the process further comprises, prior to the conversion of a starch-containing material to sugars/dextrins the steps of:

- 25 (x) reducing the particle size of the starch-containing material; and
(y) forming a slurry comprising the starch-containing material and water.

In an embodiment, the starch-containing material is milled to reduce the particle size. In an embodiment the particle size is reduced to between 0.05-3.0 mm, preferably 0.1-0.5 mm, or so that at least 30%, preferably at least 50%, more preferably at least 70%, even more
30 preferably at least 90% of the starch-containing material fits through a sieve with a 0.05-3.0 mm screen, preferably 0.1-0.5 mm screen.

The aqueous slurry may contain from 10-55 wt. % dry solids (DS), preferably 25-45 wt. % dry solids (DS), more preferably 30-40 wt. % dry solids (DS) of starch-containing material.

Conventional starch-conversion processes, such as liquefaction and saccharification
35 processes are described, *e.g.*, in U.S. Patent No. 3,912,590, EP 252730 and EP 063909, which are incorporated herein by reference.

In an embodiment, the conversion process degrading starch to lower molecular weight carbohydrate components such as sugars or fat replacers includes a debranching step.

In the case of converting starch into a sugar, the starch is depolymerized. Such a depolymerization process consists of, e.g., a pre-treatment step and two or three consecutive process steps, i.e., a liquefaction process, a saccharification process, and depending on the desired end-product, an optional isomerization process.

When the desired final sugar product is, e.g., high fructose syrup the dextrose syrup may be converted into fructose. After the saccharification process, the pH is increased to a value in the range of 6-8, e.g., pH 7.5, and the calcium is removed by ion exchange. The dextrose syrup is then converted into high fructose syrup using, e.g., an immobilized glucose isomerase.

The compositions according to the invention are useful in conventional starch processes comprising liquefaction and saccharification. In a particular aspect the invention therefore relates to a method of saccharifying a starch containing material, comprising:

(a) liquefying the starch containing material in the presence of at least an acid alpha amylase;

(b) saccharifying the starch containing material in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound, and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, and at least a glucoamylase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or the cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition.

Production of Fermentation Products

Fermentable sugars (e.g., dextrans, monosaccharides, particularly glucose) are produced from enzymatic saccharification. These fermentable sugars may be further purified and/or converted to useful sugar products. In addition, the sugars may be used as a fermentation feedstock in a microbial fermentation process for producing end-products, such as alcohol (e.g., ethanol, and butanol), organic acids (e.g., succinic acid, 3-HP and lactic acid), sugar alcohols (e.g., glycerol), ascorbic acid intermediates (e.g., gluconate, 2-keto-D-gluconate, 2,5-diketo-D-gluconate, and 2-keto-L-gulonic acid), amino acids (e.g., lysine), proteins (e.g., antibodies and fragment thereof).

In an embodiment, the fermentable sugars obtained during the liquefaction process steps are used to produce alcohol and particularly ethanol.

The organism used in fermentation will depend on the desired end-product. Typically, if ethanol is the desired end product yeast will be used as the fermenting organism. In some preferred embodiments, the ethanol-producing microorganism is a yeast and specifically *Saccharomyces* such as strains of *S. cerevisiae* (U.S. Patent No. 4,316,956). A variety of *S. cerevisiae* are commercially available and these include but are not limited to FALI (Fleischmann's Yeast), SUPERSTART (Alltech), FERMIOL (DSM Specialties), RED STAR (Lesaffre) and Angel alcohol yeast (Angel Yeast Company, China). The amount of starter yeast employed in the methods is an amount effective to produce a commercially significant amount of ethanol in a suitable amount of time, (e.g., to produce at least 10% ethanol from a substrate having between 25-40% DS in less than 72 hours). Yeast cells are generally supplied in amounts of about 10^4 to about 10^{12} , and preferably from about 10^7 to about 10^{10} viable yeast count per mL of fermentation broth. After yeast is added to the mash, it is typically subjected to fermentation for about 24-96 hours, e.g., 35-60 hours. The temperature is between about 26-34°C, typically at about 32°C, and the pH is from pH 3-6, e.g., around pH 4-5.

The fermentation may include, in addition to a fermenting microorganisms (e.g., yeast), nutrients, and additional enzymes, including phytases. The use of yeast in fermentation is well known in the art.

In further embodiments, use of appropriate fermenting microorganisms, as is known in the art, can result in fermentation end product including, e.g., glycerol, 1,3-propanediol, gluconate, 2-keto-D-gluconate, 2,5-diketo-D-gluconate, 2-keto-L-gulonic acid, succinic acid, lactic acid, amino acids, and derivatives thereof. More specifically when lactic acid is the desired end product, a *Lactobacillus* sp. (*L. casei*) may be used; when glycerol or 1,3-propanediol are the desired end-products *E. coli* may be used; and when 2-keto-D-gluconate, 2,5-diketo-D-gluconate, and 2-keto-L-gulonic acid are the desired end products, *Pantoea citrea* may be used as the fermenting microorganism. The above enumerated list are only examples and one skilled in the art will be aware of a number of fermenting microorganisms that may be used to obtain a desired end product.

The present invention also relates to methods of fermenting a retrograded starch material, comprising: fermenting the retrograded starch material with one or more (e.g., several) fermenting microorganisms, wherein the retrograded starch material is saccharified with an composition in the presence of at least an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase. The retrograded starch material may be the only substrate or it may be comprised in a starch containing material. Retrograded starch may e.g., be formed by liquefying and subsequent cooling of a starch containing material. In one aspect,

the fermenting of the retrograded starch material produces a fermentation product. In another aspect, the method further comprises recovering the fermentation product from the fermentation.

A particular aspect of the present invention relates to a method for producing a fermentation product, comprising:

(a) liquefying a starch containing material in the presence of at least an alpha amylase,

(b) saccharifying the starch containing material in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, and at least a glucoamylase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or the cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition;

(c) fermenting the saccharified starch material with one or more fermenting microorganisms to produce the fermentation product.

In one aspect, the sulfur-containing compound or the heterocyclic (reducing agent) compound is recovered following saccharification or fermentation and recycled back to a new saccharification reaction. Recycling of the sulfur-containing compound or heterocyclic compound can be accomplished using processes conventional in the art.

In one particular embodiment saccharification and fermentation is performed simultaneously.

The fermentation product may in one embodiment be recovered.

The methods of the present invention can be implemented using any conventional starch processing apparatus configured to operate in accordance with the invention.

The term "fermenting organism" refers to any organism, including bacterial and fungal organisms, such as yeast and filamentous fungi, suitable for producing a desired fermentation product. Suitable fermenting organisms are able to ferment, *i.e.*, convert, fermentable sugars, such as arabinose, fructose, glucose, maltose, mannose, or xylose, directly or indirectly into the desired fermentation product.

Examples of fermenting organisms include fungal organisms such as yeast. Preferred yeast include strains of *Saccharomyces*, in particular *Saccharomyces cerevisiae* or *Saccharomyces uvarum*; strains of *Pichia*, in particular *Pichia stipitis* such as *Pichia stipitis* CBS

5773 or *Pichia pastoris*; strains of *Candida*, in particular *Candida arabinofementans*, *Candida boidinii*, *Candida diddensii*, *Candida shehatae*, *Candida sonorensis*, *Candida tropicalis*, or *Candida utilis*. Other fermenting organisms include strains of *Hansenula*, in particular *Hansenula anomala* or *Hansenula polymorpha*; strains of *Kluyveromyces*, in particular
5 *Kluyveromyces fragilis* or *Kluyveromyces marxianus*; and strains of *Schizosaccharomyces*, in particular *Schizosaccharomyces pombe*.

Preferred bacterial fermenting organisms include strains of *Escherichia*, in particular *Escherichia coli*, strains of *Zymomonas*, in particular *Zymomonas mobilis*, strains of *Zymobacter*, in particular *Zymobacter palmae*, strains of *Klebsiella* in particular *Klebsiella*
10 *oxytoca*, strains of *Leuconostoc*, in particular *Leuconostoc mesenteroides*, strains of *Clostridium*, in particular *Clostridium butyricum*, strains of *Enterobacter*, in particular *Enterobacter aerogenes*, and strains of *Thermoanaerobacter*, in particular *Thermoanaerobacter* BG1L1 (*Appl. Microbiol. Biotech.* 77: 61-86), *Thermoanaerobacter ethanolicus*,
15 *Thermoanaerobacter mathranii*, or *Thermoanaerobacter thermosaccharolyticum*. Strains of *Lactobacillus* are also envisioned as are strains of *Corynebacterium glutamicum R*, *Bacillus thermoglucosidasius*, and *Geobacillus thermoglucosidasius*.

In an embodiment the fermenting organism is a C6 sugar fermenting organism, such as a strain of, e.g., *Saccharomyces cerevisiae*.

In one embodiment the fermenting organism is added to the fermentation medium so
20 that the viable fermenting organism, such as yeast, count per mL of fermentation medium is in the range from 10^5 to 10^{12} , preferably from 10^7 to 10^{10} , especially about 5×10^7 .

Yeast is the preferred fermenting organism for ethanol fermentation. Preferred are strains of *Saccharomyces*, especially strains of the species *Saccharomyces cerevisiae*, preferably strains which are resistant towards high levels of ethanol, i.e., up to, e.g., about 10, 12, 15 or 20 vol. %
25 or more ethanol.

Commercially available yeast include LNF SA-1, LNF BG-1, LNF PE-2, and LNF CAT-1 (available from LNF Brazil), RED STAR™ and ETHANOL RED™ yeast (available from Fermentis/Lesaffre, USA), FALI (available from Fleischmann's Yeast, USA), SUPERSTART and THERMOSACC™ fresh yeast (available from Ethanol Technology, WI, USA), BIOFERM
30 AFT and XR (available from NABC - North American Bioproducts Corporation, GA, USA), GERT STRAND (available from Gert Strand AB, Sweden), and FERMIOL (available from DSM Specialties).

The fermenting organism capable of producing a desired fermentation product from fermentable sugars is preferably grown under precise conditions at a particular growth rate.
35 When the fermenting organism is introduced into/added to the fermentation medium the inoculated fermenting organism pass through a number of stages. Initially growth does not occur. This period is referred to as the "lag phase" and may be considered a period of

adaptation. During the next phase referred to as the “exponential phase” the growth rate gradually increases. After a period of maximum growth the rate ceases and the fermenting organism enters “stationary phase”. After a further period of time the fermenting organism enters the “death phase” where the number of viable cells declines.

5 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 starch or hydrolysate and the fermentation is performed for about 8 to about 96 hours, such as about 24 to about 60 hours. The temperature is typically between about 26°C to about 60°C, in particular about 32°C
10 to 45°C, and at about pH 3 to about pH 8, such as around pH 4-5, 6, or 7.

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

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

Starch-Containing Materials

Any suitable starch-containing starting material may be used in a process of the present invention. The starting material is generally selected based on the desired fermentation
30 product. Examples of starch-containing starting materials, suitable for use in the processes of the present invention, include barley, beans, cassava, cereals, corn, milo, peas, potatoes, rice, rye, sago, sorghum, sweet potatoes, tapioca, wheat, and whole grains, or any mixture thereof. The starch-containing material may also be a waxy or non-waxy type of corn and barley. In a preferred embodiment the starch-containing material is corn. In a preferred embodiment the
35 starch-containing material is wheat.

Fermentation products

A fermentation product can be any substance derived from the fermentation. The fermentation product can be, without limitation, an alcohol (e.g., arabinitol, n-butanol, isobutanol, ethanol, glycerol, methanol, ethylene glycol, 1,3-propanediol [propylene glycol], butanediol, glycerin, sorbitol, and xylitol); an alkane (e.g., pentane, hexane, heptane, octane, nonane, decane, undecane, and dodecane), a cycloalkane (e.g., cyclopentane, cyclohexane, cycloheptane, and cyclooctane), an alkene (e.g. pentene, hexene, heptene, and octene); an amino acid (e.g., aspartic acid, glutamic acid, glycine, lysine, serine, and threonine); a gas (e.g., methane, hydrogen (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 (e.g., several) hydroxyl moieties. In a more preferred aspect, the alcohol is n-butanol. In another more preferred aspect, the alcohol is isobutanol. In another more preferred aspect, the alcohol is ethanol. In another more preferred aspect, the alcohol is methanol. In another more preferred aspect, the alcohol is arabinitol. In another more preferred aspect, the alcohol is butanediol. In another more preferred aspect, the alcohol is ethylene glycol. In another more preferred aspect, the alcohol is glycerin. In another more preferred aspect, the alcohol is glycerol. In another more preferred aspect, the alcohol is 1,3-propanediol. In another more preferred aspect, the alcohol is sorbitol. In another more preferred aspect, the alcohol is xylitol. See, for example, Gong, C. S., Cao, N. J., Du, J., and Tsao, G. T., 1999, Ethanol production from renewable resources, in *Advances in Biochemical Engineering/Biotechnology*, Scheper, T., ed., Springer-Verlag Berlin Heidelberg, Germany, 65: 207-241; Silveira and Jonas, 2002, The biotechnological production of sorbitol, *Appl. Microbiol. Biotechnol.* 59: 400-408; Nigam and Singh, 1995, Processes for fermentative production of xylitol – a sugar substitute, *Process Biochemistry* 30(2): 117-124; Ezeji *et al.*, 2003, Production of acetone, butanol and ethanol by *Clostridium beijerinckii* BA101 and *in situ* recovery by gas stripping, *World Journal of Microbiology and Biotechnology* 19(6): 595-603.

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 *et al.*, 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, 1997, Anaerobic digestion of biomass for methane production: A review, *Biomass and Bioenergy* 13(1-2): 83-114.

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 acetonc 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 and Lee, 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 numbered paragraphs:

Paragraph [1]. A composition comprising: (a) an X143 polypeptide having retrograded starch degrading activity and (b) a sulfur-containing compound, and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the X143 polypeptide having retrograded starch degrading activity and the sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, enhances degradation of retrograded starch material by the X143 polypeptide.

Paragraph [2]. The composition of claim 1, wherein the X143 polypeptide is selected from the group consisting of: a polypeptide having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity to the mature polypeptide of SEQ ID NO: 2.

Paragraph [3]. The composition of paragraph 1 or 2, wherein the sulfur-containing compound is selected from the group consisting of: (I-1): ethanethiol; (I-2): 2-propanethiol; (I-3): 2-propene-1-thiol; (I-4): 2-mercaptoethanesulfonic acid; (I-6): benzenethiol; (I-7): benzene-1,2-dithiol; (I-5): cysteine; (II-1): methionine; (II-2): glutathione; (II-3): cystine; or a salt or solvate thereof.

Paragraph [4]. The composition of paragraph 3, wherein the sulfur-containing compound is L-cysteine.

Paragraph [5]. The composition of paragraph 1 or 2, wherein the heterocyclic compound is selected from the group consisting of: (I-1): (1,2-dihydroxyethyl)-3,4-dihydroxyfuran-2(5H)-one; (I-2): 4-hydroxy-5-methyl-3-furanone; (I-3): 5-hydroxy-2(5H)-furanone; (I-4): [1,2-dihydroxyethyl]furan-2,3,4(5H)-trione; (I-5): α -hydroxy- γ -butyrolactone; (I-6): ribonic γ -lactone; (I-7): glucuronic acid γ -lactone; (I-8): dihydrobenzofuran; (I-9): 5-(hydroxymethyl)furfural; (I-10): furoin; (I-11): 2(5H)-furanone; (II-1): gluconic acid δ -lactone; (II-2): 4-hydroxycoumarin; (II-3): 5,6-dihydro-2H-pyran-2-one; (II-4): 5,6-dihydro-4-hydroxy-6-methyl-2H-pyran-2-one; (II-5): 1,5-anhydro-2-deoxy-arabino-hex-1-enitol; and (II-6): 3-deoxy-erythro-hexosulose; 3-hydroxy-5-methylisoxazole; or a salt or solvate thereof.

Paragraph [6]. The composition of paragraph 5, wherein the heterocyclic compound is (1,2-dihydroxyethyl)-3,4-dihydroxyfuran-2(5H)-one.

Paragraph [7]. The composition of paragraph 1 or 2, wherein the sulfur-containing compound is a sulfur(IV) oxyanion, particularly sodium bisulfite (NaHSO_3).

Paragraph [8]. The composition of paragraph 1 or 2, wherein the cellobiose dehydrogenase is a *Humicola insolens* cellobiose dehydrogenase.

Paragraph [9]. The composition of any of paragraphs 1-8, which composition further comprises (c) one or more (e.g., several) polypeptides selected from the group consisting of α -amylase, a glucoamylase, a pullulanase, a protease, a phytase, and optionally one or more 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, and a swollenin.

Paragraph [10]. The composition of paragraph 9, wherein the cellulases are selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase.

Paragraph [11]. A method for degrading a retrograded starch material, comprising: treating the retrograded starch material with a composition in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition.

Paragraph [12]. The method of paragraph 11, wherein the retrograded starch material is comprised in a starch containing material, and wherein the starch containing material is selected from the group consisting of barley, beans, cassava, cereals, corn, milo, peas, potatoes, rice, rye, sago, sorghum, sweet potatoes, tapioca, wheat, and whole grains, or any mixture thereof.

Paragraph [13]. The method of any of paragraphs 11-12, further comprising recovering the degraded retrograded starch material.

Paragraph [14]. The method of any of paragraphs 11-13, wherein the composition comprises one or more (*e.g.*, several) polypeptides selected from the group consisting of an acid alpha-amylase, a glucoamylase, a pullulanase, a protease, a phytase, and optionally one or more 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, and a swollenin.

Paragraph [15]. A method of saccharifying a starch containing material, comprising:

(a) liquefying the starch containing material in the presence of at least an alpha amylase;

(b) saccharifying the starch containing material in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound, and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, and at least a glucoamylase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the

mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or the cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition.

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

(a) liquefying a starch containing material in the presence of at least an alpha amylase,

(b) saccharifying the starch containing material in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, and at least a glucoamylase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or the cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition;

(c) fermenting the saccharified starch material with one or more fermenting microorganisms to produce the fermentation product.

Paragraph [17]. The method of paragraph 15 or 16, wherein the methods are performed in the presence of further enzymes comprising one or more enzymes selected from the group consisting of alpha-amylase, a glucoamylase, a pullulanase, a protease, a phytase, and optionally one or more 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, and a swollenin.

Paragraph [18]. The method of paragraph 16 or 17, wherein saccharification and fermentation is performed simultaneously.

Paragraph [19]. The method of any of paragraphs 16-18, wherein the fermentation product is an alcohol, particularly ethanol.

Paragraph [20]. The method of any of paragraphs 16-19, further comprising recovering the fermentation product from the fermentation.

Paragraph [21]. The methods of any of paragraphs 11-20, wherein the X143 polypeptide is selected from the group consisting of: a polypeptide having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity to the mature polypeptide of SEQ ID NO: 2.

Paragraph [22]. The method of any of paragraphs 11-21, wherein the sulfur-containing compound is selected from the group consisting of: (I-1): ethanethiol; (I-2): 2-propanethiol; (I-3): 2-propene-1-thiol; (I-4): 2-mercaptoethanesulfonic acid; (I-6): benzenethiol; (I-7): benzene-1,2-dithiol; (I-5): cysteine; (II-1): methionine; (II-2): glutathione; (II-3): cystine; or a salt or solvate thereof.

Paragraph [23]. The method of paragraph 22, wherein the sulfur-containing compound is L-cysteine.

Paragraph [24]. The method of paragraph 23, wherein the concentration of L-cysteine is in the range from 0.5 to 3 mM, more particularly from 1 to 2 mM.

Paragraph [25]. The method of any of paragraphs 11-21, wherein the heterocyclic compound is selected from the group consisting of: (I-1): (1,2-dihydroxyethyl)-3,4-dihydroxyfuran-2(5H)-one; (I-2): 4-hydroxy-5-methyl-3-furanone; (I-3): 5-hydroxy-2(5H)-furanone; (I-4): [1,2-dihydroxyethyl]furan-2,3,4(5H)-trione; (I-5): α -hydroxy- γ -butyrolactone; (I-6): ribonic γ -lactone; (I-7): glucuronic acid γ -lactone; (I-8): dihydrobenzofuran; (I-9): 5-(hydroxymethyl)furfural; (I-10): furoin; (I-11): 2(5H)-furanone; (II-1): gluconic acid δ -lactone; (II-2): 4-hydroxycoumarin; (II-3): 5,6-dihydro-2H-pyran-2-one; (II-4): 5,6-dihydro-4-hydroxy-6-methyl-2H-pyran-2-one; (II-5): 1,5-anhydro-2-deoxy-arabino-hex-1-enitol; and (II-6): 3-deoxy-erythro-hexosulose; 3-hydroxy-5-methylisoxazole; or a salt or solvate thereof.

Paragraph [26]. The method of paragraph 25, wherein the heterocyclic compound is (1,2-dihydroxyethyl)-3,4-dihydroxyfuran-2(5H)-one.

Paragraph [27]. The method of any of paragraphs 11-21, wherein the sulfur-containing compound is a sulfur(IV) oxyanion, particularly sodium bisulfite (NaHSO_3).

Paragraph [28]. The method of any of paragraphs 11-21, wherein the cellobiose dehydrogenase is a *Humicola insolens* cellobiose dehydrogenase.

Paragraph [29]. The method of any of paragraphs 15-28, wherein the starch containing material, is selected from the group consisting of barley, beans, cassava, cereals, corn, milo, peas, potatoes, rice, rye, sago, sorghum, sweet potatoes, tapioca, wheat, and whole grains, or any mixture thereof.

5

Paragraph [30]. The method of any of paragraphs 14 and 17-29, wherein the cellulases are selected from the group consisting of an endoglucanase, a cellobiohydrolase, and a beta-glucosidase.

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

Examples

15 **Example 1: Methods of evaluating the effect of sulfur-containing or reducing agent compound on X143 polypeptide having retrograded starch degrading activity**

The effect of sulfur-containing or reducing agent compounds on the retrograded starch hydrolyzing activity of X143 polypeptide was evaluated according to the procedures described below.

20 Retrograded starch was prepared from pure corn starch purchased from Sigma-Aldrich (St. Louis, MO, USA). 26.9 g of pure corn starch was mixed with 473.1 g of MilliQ water in total 500 g (approximately 5% dry solid) and transferred into a 1 liter size Scott Bottle with blue cap. Starch gelatinization was performed at 85°C for 75 min in a temperature-control water-bath. During the gelatinization process, the starch slurry in bottle was mixed every 15 min interval to eliminate starch clump and to provide more uniform gelatinization process. After completion of 25 75 min gelatinization, the gelatinized starch in bottle was cooled to room temperature. The dry solid of the gelatinized starch was determined to be approximately 5%. Sodium azide (0.025%) was added to the gelatinized starch to prevent microbial growth and then homogenized with a homogenizer for 1 min.

30 Starch retrogradation process was carried out by performing multiple freezing and thawing cycles as previously reported by *Karim et al* and *Szymonska et al*. The gelatinized starch was freeze overnight at -20°C. After overnight, the frozen starch was slowly thawed at room temperature with occasional mixing. After completely thawed, the starch was re-freezing again at -20°C for overnight and then followed by slow thawing of the freeze starch at room temperature. The freeze and thaw cycles were repeated for 5 days. After multiple freeze and 35 thaw steps, the resulted retrograded starch will be used as substrate for X143 polypeptide activity assay.

The degradation of retrograded starch was conducted using 2.0 ml deep-well plates (Axygen Scientific, Union City, CA, USA) in a total reaction volume of 1.0 ml. Each degradation was performed with retrograded starch (1% dry solid) of 50 mM sodium acetate pH 5.0 buffer containing 20 μ M copper sulfate. For the degradation of retrograded starch, *Aspergillus* 5 *nidulans* X143 polypeptide was dosed at 25 μ g protein per ml with and without sulfur-containing or reducing agent compound at a specified concentration and 100 μ g protein per ml of alpha-glucosidase. For control, the degradation of retrograded starch, the alpha-glucosidase was dosed at 100 μ g protein per ml with and without sulfur-containing or reducing agent compound at a specified concentration and without the X143 polypeptide. Alpha-glucosidase was added 10 solely to generate glucose from oligosaccharides that produce from any degrading activity of X143 polypeptide. The plate was then sealed using an ALPS-300™ or ALPS-3000™ plate heat sealer (Abgene, Epsom, United Kingdom), mixed thoroughly, and incubated at 40°C for 2 hours with 800 rpm shaking in a Thermomixer comfort (Eppendorf AG, Hamburg, Germany). All experiments were performed at least in triplicate.

Following degradation, samples in well plates were subjected to centrifugation at 2000 15 rpm for 5 min to spin down the starch and water vapor. The seal was removed and the glucose concentration from the supernatant was determined with WAKO Autokit Glucose assay kits (Wako Pure Chemical, Ltd, Osaka, Japan). 10 μ l of reaction mixture from the supernatant was transferred to a 96-wells plate followed by adding 200 μ l of Autokit Glucose developing solution. 20 The plate was stand for 20 min at room temperature and then measure absorbance at 505 nm. Glucose was quantified based on the increase in absorbance at 505 nm with reference to the standard curve of glucose prepared using the Autokit Glucose. The resultant glucose was used to calculate the percentage boosting of activity for each reaction. Percentage boosting of activity was calculated from the ratio of net concentration of X143 polypeptide-produced 25 glucose subtracting the background glucose concentration from a control in which no X143 polypeptide was added. Data were processed using MICROSOFT EXCEL™ software (Microsoft, Richland, WA, USA).

The sulfur-containing compound evaluated includes L-cysteine and the reducing agent compound evaluated includes L-ascorbic acid. The compounds were obtained from Sigma- 30 Aldrich Co. (St. Louis, Missouri, USA).

Example 2: Effect of L-cysteine or L-ascorbic acid on *Aspergillus nidulans* X143 polypeptide degradation of retrograded starch

The effect of L-cysteine or L-ascorbic acid on the degrading activity of the *A. nidulans* 35 X143 polypeptide on retrograded starch was determined using the experimental conditions and procedures described in Example 1. The concentration of L-cysteine or L-ascorbic acid was 1 or 4 mM, respectively.

As shown in Table 1, without the presence of L-Cysteine or L-Ascorbic acid, *A. nidulans* X143 polypeptide did not show any significant activity toward retrograded starch. For controls, alpha-glucosidase alone without X143 and with L-Cysteine or L-Ascorbic acid, there was no degrading activity on retrograded starch. Only with the presence of L-Cysteine or L-ascorbic acid, the degradation activity of X143 polypeptide on retrograded starch increase significantly, particularly with L-Cysteine. These results demonstrated that compound such as L-Cysteine or L-Ascorbic acid is require for the enzymatic reaction of *A. nidulans* X143 polypeptide toward retrograded starch.

Table 1

Compound	% Boosting	
	Without X143	With X143
No compound	0.0%	4.8%
L-Cysteine	0.0%	102.7%
L-Ascorbic acid	0.0%	28.9%

Example 3: Effect of L-cysteine concentration on *Aspergillus nidulans* X143 polypeptide on degradation of retrograded starch

The effect of L-cysteine at various concentrations on the boosting activity of the *A. nidulans* X143 polypeptide during degradation of retrograded starch was determined using the experimental conditions and procedures described in Example 1. The concentration of L-cysteine was 0, 0.5, 1, 2 or 4 mM.

As shown in Table 2, without the presence of L-Cysteine, *A. nidulans* X143 polypeptide did not show any significant activity. Hydrolysis activity of X143 polypeptide on retrograded starch increases with the presence of L-Cysteine, and activity increase even greater with increased concentration of L-Cysteine and reached an optimal point corresponding to the concentration of L-Cysteine added.

Table 2

L-Cysteine (mM)	% Boosting	
	Without X143	With X143
0	0.0%	3.8%
0.5	0.0%	75.8%
1	0.0%	94.5%
2	0.0%	123.0%
3	0.0%	64.9%
4	0.0%	57.7%

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.

Claims

1. A composition comprising: (a) an X143 polypeptide having retrograded starch degrading activity and (b) a sulfur-containing compound, and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the X143 polypeptide having retrograded starch degrading activity and the sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, enhances degradation of retrograded starch material by the X143 polypeptide.
2. The composition of claim 1, wherein the X143 polypeptide is selected from the group consisting of: a polypeptide having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity to the mature polypeptide of SEQ ID NO: 2.
3. The composition of claim 2, wherein the sulfur-containing compound is L-cysteine.
4. The composition of claim 1, wherein the heterocyclic compound is (1,2-dihydroxyethyl)-3,4-dihydroxyfuran-2(5H)-one.
5. The composition of claim 1, wherein the sulfur-containing compound is a sulfur(IV) oxyanion, particularly sodium bisulfite (NaHSO_3).
6. The composition of claim 1, wherein the cellobiose dehydrogenase is a *Humicola insolens* cellobiose dehydrogenase.
7. A method for degrading a retrograded starch material, comprising: treating the retrograded starch material with a composition comprising an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and wherein the combination of the polypeptide having

retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition.

- 5 8. The method of claim 7, wherein the retrograded starch material is comprised in a starch containing material, and wherein the starch containing material is selected from the group consisting of barley, beans, cassava, cereals, corn, milo, peas, potatoes, rice, rye, sago, sorghum, sweet potatoes, tapioca, wheat, and whole grains, or any mixture thereof.
- 10 9. A method of saccharifying a starch containing material, comprising:
- (a) liquefying the starch containing material in the presence of at least an alpha amylase;
- (b) saccharifying the starch containing material in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound, and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, and at least a glucoamylase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and
- 15 and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, and at least a glucoamylase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and
- 20 wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or the cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition.
- 25 10. A method for producing a fermentation product, comprising:
- (a) liquefying a starch containing material in the presence of at least an alpha amylase,
- (b) saccharifying the starch containing material in the presence of an X143 polypeptide having retrograded starch degrading activity and a sulfur-containing compound and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, and at least a glucoamylase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and
- 30 and/or a heterocyclic compound (reducing agent), and/or a cellobiose dehydrogenase, and at least a glucoamylase, wherein the X143 polypeptide is selected from the group consisting of: (i) a polypeptide having at least 75% sequence identity to the mature polypeptide of SEQ ID NO: 2; (ii) a polypeptide encoded by a polynucleotide having at least 75% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 or the cDNA sequence thereof; and
- 35 wherein the combination of the polypeptide having retrograded starch degrading activity and the sulfur-containing compound, and/or the heterocyclic compound (reducing agent), and/or the

cellobiose dehydrogenase, enhances degradation of the retrograded starch material by the enzyme composition;

(c) fermenting the saccharified starch material with one or more fermenting microorganisms to produce the fermentation product.

5

11. The method of claim 10, wherein the fermentation product is an alcohol, particularly ethanol.

10 12. The methods of any of claims 7-11, wherein the X143 polypeptide is selected from the group consisting of: a polypeptide having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity to the mature polypeptide of SEQ ID NO: 2.

15 13. The method of any of claims 7-12, wherein the sulfur-containing compound is L-cysteine.

14. The method of any of claims 7-12, wherein the heterocyclic compound is (1,2-dihydroxyethyl)-3,4-dihydrofuran-2(5H)-one.

20 15. The method of any of claims 7-12, wherein the sulfur-containing compound is a reducing agent and is sulfur(IV) oxyanion, particularly sodium bisulfite (NaHSO₃).

25 16. The method of any of claims 7-12, wherein the cellobiose dehydrogenase is a *Humicola insolens* cellobiose dehydrogenase.

17. The method of any of claims 9-16, wherein the starch containing material, is selected from the group consisting of barley, beans, cassava, cereals, corn, milo, peas, potatoes, rice, rye, sago, sorghum, sweet potatoes, tapioca, wheat, and whole grains, or any mixture thereof.

30

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/041113**A. CLASSIFICATION OF SUBJECT MATTER****C12N 9/22(2006.01)i, C12P 7/06(2006.01)i**

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N 9/22; C12N 15/63; C12P 19/20; C13J 1/06; C12P 7/14; C12N 15/31; C12P 7/06

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

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS(KIPO internal) & Keywords: X143 polypeptide, sulfur-containing compound, retrograded starch, saccharifying, fermenting

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2010-059413 A2 (NOVOZYMES, INC.) 27 May 2010 See abstract; claims 1-3; and sequence list.	1-12
A	NCBI, Genbank accession no. XP_663067.1 (09 April 2008) See the whole sequence.	1-12
A	US 2008-0121227 A1 (BHARGAVA, SWAPNIL et al.) 29 May 2008 See the whole document.	1-12
A	WO 87-04465 A1 (BJURENVALL, INGEMAR et al.) 30 July 1987 See the whole document.	1-12
A	WO 2012-109119 A2 (NOVOZYMES NORTH AMERICA, INC.) 16 August 2012 See the whole document.	1-12



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) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

26 September 2014 (26.09.2014)

Date of mailing of the international search report

26 September 2014 (26.09.2014)

Name and mailing address of the ISA/KR

International Application Division
Korean Intellectual Property Office
189 Cheongsu-ro, Seo-gu, Daejeon Metropolitan City, 302-701,
Republic of Korea

Facsimile No. +82-42-472-7140

Authorized officer

HEO, Joo Hyung

Telephone No. +82-42-481-8150



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2014/041113**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.: 13-17
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:

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 any 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.:

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.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2014/041113

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2010-059413 A2	27/05/2010	CN 102292444 A EP 2358878 A2 US 2011-0283421 A1 WO 2010-059413 A3	21/12/2011 24/08/2011 17/11/2011 19/08/2010
US 2008-0121227 A1	29/05/2008	WO 2006-028897 A2 WO 2006-028897 A3	16/03/2006 21/09/2006
WO 87-04465 A1	30/07/1987	AU 6899687 A	14/08/1987
WO 2012-109119 A2	16/08/2012	EP 2673368 A2 US 2013-0316407 A1 WO 2012-109119 A3	18/12/2013 28/11/2013 28/02/2013