

METHOD FOR REDUCING VISCOSITY IN SACCHARIFICATION PROCESS**CROSS REFERENCE TO RELATED APPLICATIONS**

[0001] This application claims the benefit of U.S. Provisional Application No. 61/453,923, filed March 17, 2011, which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0002] The present invention relates to compositions useful for hydrolyzing biomass, methods of using such compositions to hydrolyze biomass materials, and methods for reducing viscosity of biomass saccharification mixtures.

BACKGROUND OF THE INVENTION

[0003] Bioconversion of renewable lignocellulosic biomass to a fermentable sugar that is subsequently fermented to produce alcohol (*e.g.*, ethanol) as an alternative to liquid fuels has attracted the intensive attention of researchers since the 1970s, when the oil crisis occurred (Bungay, H. R., "Energy: the biomass options". NY: Wiley; 1981; Olsson L, Hahn-Hagerdal B. *Enzyme Microb Technol* 1996,18:312-31; Zaldivar, J et al., *Appl Microbiol Biotechnol* 2001, 56: 17-34; Galbe, M et al., *Appl Microbiol Biotechnol* 2002, 59:618-28). The production of sugars from lignocellulosic biomass materials has been known for some time, as has the subsequent fermentation and distillation of the sugars into ethanol. Much of the prior development occurred around the time of World War II when fuels were at a premium in such countries as Germany, Japan and the Soviet Union. These early processes were primarily directed to acid hydrolysis, which were complex in engineering and design, and were typically sensitive to small variations in the processes, such as to temperature, pressure and/or acid concentrations. A comprehensive discussion of these early processes is found in "Production of Sugars from Wood Using High-pressure Hydrogen Chloride", *Biotechnology and Bioengineering*, Volume XXV, at 2757-2773 (1983).

[0004] The abundant supply of petroleum in the period from World War II through the early 1970s slowed ethanol conversion research. However, due to the oil crisis of 1973, researchers increased their efforts to develop processes for the utilization of wood and agricultural byproducts for the production of ethanol. This research was especially important for development of ethanol as a gasoline additive to reduce the dependency of the United States upon foreign oil production, to increase the octane rating of fuels, and to reduce exhaust pollutants as an environmental measure.

[0005] Concurrently with the "oil crisis," the U.S. Environmental Protection Agency

promulgated regulations requiring reduced lead additives. Insofar as ethanol is virtually a replacement of lead, some refineries have selected ethanol as the substitute for its capability of easy introduction into a refinery's operation without costly capital equipment investment.

[0006] The high pressure and high temperature gas saccharification processes developed decades ago continue to be improved. New and current research focuses greatly on enzymatic conversion processes, which employ enzymes from a variety of organisms, such as mesophilic and thermophilic fungi, yeast and bacteria, degrading cellulose into fermentable sugars. Uncertainty remains with these processes, mainly on their ability to be scaled up for commercialization and on the efficiency of ethanol production.

[0007] Cellulose and hemicellulose are the most abundant plant materials produced by photosynthesis. They can be degraded for use as an energy source by numerous microorganisms, including bacteria, yeast and fungi, which produce enzymes capable of hydrolysis of the polymeric substrates to monomeric sugars (Aro et al., 2001). Organisms are often restrictive with regard to which sugars they use, and this dictates which sugars are best to produce during conversion. As we approach the limits of non-renewable resources, we recognize the enormous potential of cellulose to become a major renewable energy resource (Krishna et al., 2001). The effective utilization of cellulose through biological processes can potentially overcome the shortage of foods, feeds, and fuels (Ohmiya et al., 1997).

[0008] Cellulases are enzymes that hydrolyze cellulose (beta-1,4-glucan or beta D-glucosidic linkages) resulting in the formation of glucose, cellobiose, cellooligosaccharides, and the like. Cellulases have been traditionally divided into 3 major classes: endoglucanases (EC 3.2.1.4) ("EG"), exoglucanases or cellobiohydrolases (EC 3.2.1.91) ("CBH") and beta-glucosidases ([beta] -D-glucoside glucohydrolase; EC 3.2.1.21) ("BG") (Knowles et al., 1987 and Shulein, 1988). Endoglucanases act mainly on the amorphous parts of the cellulose fiber, whereas cellobiohydrolases are also able to degrade crystalline cellulose.

[0009] Cellulases have also been shown to be useful in degradation of cellulose biomass to ethanol (wherein the cellulases degrade cellulose to glucose, and yeast or other microbes further ferment the glucose into ethanol), in the treatment of mechanical pulp (Pere et al., 1996), for use as a feed additive (WO 91/04673) and in grain wet milling. Separate saccharification and fermentation is a process whereby cellulose present in biomass, *e.g.*, corn stover, is converted to glucose and subsequently yeast strains convert glucose into ethanol. Simultaneous saccharification and fermentation is a process whereby cellulose

present in biomass, *e.g.*, corn stover, is converted to glucose and, at the same time and in the same reactor, yeast strains convert glucose into ethanol. Ethanol production from readily available sources of cellulose provides a stable, renewable fuel source.

[0010] Cellulases are produced by a number of bacteria, yeast and fungi. Certain fungi produce a complete cellulase system (*i.e.*, a whole cellulase) capable of degrading crystalline forms of cellulose. A whole cellulase, especially one that is naturally occurring, is, however, not necessarily capable of achieving efficient degradation because it may not include all the components/activities required for this efficiency, for example, activities from each of the CBH, EG and BG classifications. (Filho et al., 1996). It is known that individual CBH, EG, and BG components alone do not bring about efficient hydrolysis, but the combination of EG-type cellulases and CBH- type cellulases interact to more efficiently degrade cellulose than either enzyme used alone (Wood, 1985; Baker et al., 1994; and Nieves et al., 1995).

[0011] Cellulases are known in the art to be useful in the treatment of textiles, for enhancing the cleaning ability of detergent compositions, for use as a softening agent, for improving the feel and appearance of cotton fabrics, and the like (Kumar et al., 1997). Cellulase-containing detergent compositions with improved cleaning performance (US Pat. No. 4,435,307; GB App. Nos. 2,095,275 and 2,094,826) and for use in the treatment of fabric to improve the feel and appearance of the textile (US Pat. Nos. 5,648,263, 5,691,178, and 5,776,757, and GB App. No. 1,358,599), have been described.

[0012] Hence, cellulases produced in fungi and bacteria have received significant attention. In particular, fermentation of *Trichoderma spp.* (*e.g.*, *T. longibrachiatum* or *T. reesei*) has been shown to produce a complete cellulase system capable of degrading crystalline forms of cellulose. Over the years, *Trichoderma* cellulase production has been improved by classical mutagenesis, screening, selection and development of highly refined, large scale inexpensive fermentation conditions. While the multi-component cellulase system of *Trichoderma spp.* is able to hydrolyze cellulose to glucose, there are cellulases from other microorganisms, particularly bacterial strains, with different properties for efficient cellulose hydrolysis, and it would be advantageous to express these proteins in a filamentous fungus for industrial scale cellulase production. However, the results of many studies demonstrate that the yield of expressing bacterial enzymes from filamentous fungi is low (Jeeves et al., 1991).

[0013] Soluble sugars such as glucose and cellobiose have many uses for the production of chemicals and biological products. The optimization of cellulose hydrolysis allows for the

use of less enzymes and improved cost effectiveness for the production of soluble sugars.

[0014] An efficient conversion of lignocellulosic biomass into fermentable sugars is key to producing bioethanol in a cost-effective and environmentally-friendly way. To reduce energy and processing cost, particularly for distillation, the minimum ethanol concentration produced by a viable process should be at least 4% (w/v). Such an increased ethanol concentration can be achieved by processing substrates having high dry matter of solids. However a common problem associated with saccharifying a high dry matter biomass is the high viscosity of the slurry, resulting in a slurry that is not pumpable or requires large energy input during handling. When dealing with handling of high solids, problems such as 1) insufficient mixing with limited mass transfer, 2) increasing concentration of inhibitors, such as acetic acid, furfural, 5-hydroxymethyl furfural, phenolic lignin degradation, 3) production inhibition, such as glucose, cellobiose, ethanol, and 4) fermentation microorganism viability, will occur. High viscosity limits the dry substance level in the process, increasing energy and water consumption, reducing the separation efficiency, evaporation and heat exchange, and ultimately, the ethanol yield. Reduction of viscosity is therefore beneficial, and enzymes play a key role in breaking down the soluble/insoluble compounds causing high viscosity.

[0015] Studies to increase solid loading and/or reduce viscosity of saccharification processes have taken place. For example, a number of studies utilized fed-batch operations in order to increase the solids level in the biomass substrate loading. A gravimetric mixing reactor design was used, which allowed batch enzymatic liquefaction and hydrolysis of pretreated wheat straw at up to 40% solids concentration. This fed-batch strategy sequentially loads the biomass substrate or substrate plus enzymes during enzymatic hydrolysis in order to achieve hydrolysis of a large amount of substrate, a relatively low viscosity during hydrolysis, and a relatively high glucose concentration during the process. Alternatively, enzymatic pre-hydrolysis of a lignocellulosic biomass for a period of time at the enzymes' optimum temperature, e.g., 50°C, can be carried out to reduce the viscosity of the slurry, enabling pumping and stirring. The decrease in viscosity during pre-hydrolysis makes the subsequent fermentation or SSF possible.

[0016] Despite the development of numerous approaches, there remains a need in the art for additional ways to reduce viscosity and improve yield of desirable fermentable sugars.

[0017] All references cited herein, including patents, patent applications, and publications, are incorporated by reference in their entirety.

SUMMARY OF INVENTION

[0018] The present disclosure is based, in part, on the surprising discovery that inclusion of a certain endoglucanase enzyme (e.g., a polypeptide having glycosyl hydrolase family 61 (“GH61”)/endoglucanase activity, such as the *T. reesei* endoglucanase (“Eg4”)) in a biomass saccharification mixture substantially reduces the viscosity of the mixture. The disclosure also pertains to the inclusion of such enzyme(s) to substantially improve the saccharification and the yields of desirable fermentable sugars from a given biomass substrate.

[0019] Provided herein are polypeptides having glycosyl hydrolase family 61 (“GH61”)/endoglucanase activity. By “GH61/endoglucanase activity” it is meant that the polypeptide has a GH61 activity and/or an endoglucanase activity. In some aspects, the polypeptide is isolated. In some aspects, the polypeptide having GH61/endoglucanase activity (e.g., an isolated polypeptide) is a GH61 endoglucanase or an endoglucanase IV (“EG IV”) from various species, or a polypeptide corresponding to (e.g., sharing homology with, sharing functional domains, sharing GH61 motif(s), and/or sharing conservative residues with) a GH61 endoglucanase (e.g., a *T. reesei* Eg4 polypeptide). Such species include *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, *Chrysosporium*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Neurospora intermedia*, *Penicillium purpurogenum*, *Penicillium canescens*, *Penicillium solitum*, *Penicillium funiculosum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma*

harzianum, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, *Trichoderma viride*, *Geosmithia emersonii*, or *G. stearothermophilus*.

[0020] In some aspects, the polypeptide having GH61/endoglucanase activity (*e.g.*, an isolated polypeptide) is a GH61 endoglucanase selected from the group consisting of the

polypeptides with amino acid sequences shown in FIG. 1 of the present disclosure. For example, suitable GH61 endoglucanases include those that are represented by their GenBank Accession Numbers CAB97283.2, CAD70347.1, CAD21296.1, CAE81966.1, CAF05857.1, EAA26873.1, EAA29132.1, EAA30263.1, EAA33178.1, EAA33408.1, EAA34466.1, EAA36362.1, EAA29018.1, and EAA29347.1, or those that are named St61

from *S. thermophilum* 24630, St61A from *S. thermophilum* 23839c, St61B from *S. thermophilum* 46583, St61D from *S. thermophilum* 80312, Afu61a from *A. fumigatus* Afu3g03870 (NCBI Ref: XP_748707), an endoglucanase of NCBI Ref: XP_750843.1 from *A. fumigatus* Afu6g09540, an endoglucanase of *A. fumigatus* EDP47167, an endoglucanase of *T. terrestris* 16380, an endoglucanase of *T. terrestris* 155418, an endoglucanase of

T. terrestris 68900, Cg61A (EAQ86340.1) from *C. globosum*, *T. reesei* Eg7, *T. reesei* Eg4, and an endoglucanase with GenBank Accession: XP_752040 from *A. fumigatus* Af293. In some aspects, the polypeptide having GH61/endoglucanase activity (*e.g.*, isolated polypeptide) comprises an amino acid sequence that is at least about 60% (*e.g.*, at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%)

sequence identity to any one of SEQ ID NOs: 1-29 and 148. In certain aspects, the polypeptide having GH61/endoglucanase activity (*e.g.*, isolated polypeptide) comprises an amino acid sequence that comprises one or more sequence motif(s) selected from the group consisting of: (1) SEQ ID NOs:84 and 88; (2) SEQ ID NOs:85 and 88; (3) SEQ ID NO:86; (4) SEQ ID NO:87; (5) SEQ ID NOs:84, 88 and 89; (6) SEQ ID NOs:85, 88, and 89; (7) SEQ ID NOs: 84, 88, and 90; (8) SEQ ID NOs: 85, 88 and 90; (9) SEQ ID NOs:84, 88 and 91; (10) SEQ ID NOs: 85, 88 and 91; (11) SEQ ID NOs: 84, 88, 89 and 91; (12) SEQ ID NOs: 84, 88, 90 and 91; (13) SEQ ID NOs: 85, 88, 89 and 91; and (14) SEQ ID NOs: 85, 88, 90 and 91. In some embodiments, the polypeptide is at least about 100 (*e.g.*, at least about 120, 130, 140, 150, 160, 170, 180, 190, 200, 220, 240, or more) amino acid residues in length.

[0021] In some aspects, the polypeptide having GH61/endoglucanase activity is a variant of a GH61 endoglucanase such as, for example, one selected from those listed in FIG. 1.

Suitable polypeptide include, *e.g.* GenBank Accession Number CAB97283.2, CAD70347.1,

CAD21296.1, CAE81966.1, CAF05857.1, EAA26873.1, EAA29132.1, EAA30263.1, EAA33178.1, EAA33408.1, EAA34466.1, EAA36362.1, EAA29018.1, or EAA29347.1, or St61 of *S. thermophilum* 24630, St61A of *S. thermophilum* 23839c, St61B of *S.*

thermophilum 46583, St61D of *S. thermophilum* 80312, Afu61a of *A. fumigatus* Afu3g03870

5 (NCBI Ref: XP_748707), an enzyme of *A. fumigatus* Afu6g09540 (NCBI Ref: XP_750843.1), an enzyme of *A. fumigatus* EDP47167, an enzyme of *T. terrestris* 16380, an enzyme of *T. terrestris* 155418, an enzyme of *T. terrestris* 68900, and *C. globosum* Cg61A (EAQ86340.1), *T. reesei* Eg7, *T. reesei* Eg4, and an enzyme of *A. fumigatus* Af293 (with GenBank Accession: XP_752040). In some aspects, the polypeptide having GH61/

10 endoglucanase activity is a variant of an enzyme comprising any one of SEQ ID NOs: 1-29 and 148. The polypeptide having GH61/endoglucanase activity may be a variant of an enzyme having at least about 100 (e.g., at least about 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 220, 240 or more) amino acid residues in length, comprising one or more of the sequence motifs selected from: (1) SEQ ID NOs:84 and 88; (2) SEQ ID NOs:85 and 88; (3) 15 SEQ ID NO:86; (4) SEQ ID NO:87; (5) SEQ ID NOs:84, 88 and 89; (6) SEQ ID NOs:85, 88, and 89; (7) SEQ ID NOs: 84, 88, and 90; (8) SEQ ID NOs: 85, 88 and 90; (9) SEQ ID NOs:84, 88 and 91; (10) SEQ ID NOs: 85, 88 and 91; (11) SEQ ID NOs: 84, 88, 89 and 91; (12) SEQ ID NOs: 84, 88, 90 and 91; (13) SEQ ID NOs: 85, 88, 89 and 91: and (14) SEQ ID NOs: 85, 88, 90 and 91 . The polypeptide having GH61/endoglucanase activity may be a 20 variant of a GH61 endoglucanase, wherein the variant has an amino acid sequence having at least about 60% (e.g., at least about any of 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%) identity to any one of SEQ ID NOs:1-18.

[0022] In some aspects, the polypeptide having GH61/endoglucanase activity (e.g., an isolated polypeptide, including a variant of GH61 endoglucanase) has endoglucanase activity.

25 The variant may comprise at least one motif (at least 1, 2, 3, 4, 5, 6, 7, or 8 motifs) selected from SEQ ID NOs:84-91. For the purpose of the present disclosure enzymes can be referred to by their functionalities. For example, an endoglucanase polypeptide can also be referred as polypeptide having endoglucanase activity, or vice versa.

[0023] In some aspects, the polypeptide having GH61/endoglucanase activity (including a 30 variant of GH61 endoglucanase) comprises one or more sequence motif(s) selected from: (1) SEQ ID NOs:84 and 88; (2) SEQ ID NOs:85 and 88; (3) SEQ ID NO:86; (4) SEQ ID NO:87; (5) SEQ ID NOs:84, 88 and 89; (6) SEQ ID NOs:85, 88, and 89; (7) SEQ ID NOs: 84, 88,

and 90; (8) SEQ ID NOs: 85, 88 and 90; (9) SEQ ID NOs: 84, 88 and 91; (10) SEQ ID NOs: 85, 88 and 91; (11) SEQ ID NOs: 84, 88, 89 and 91; (12) SEQ ID NOs: 84, 88, 90 and 91; (13) SEQ ID NOs: 85, 88, 89 and 91; and (14) SEQ ID NOs: 85, 88, 90 and 91.

[0024] In some aspects, the polypeptide having GH61/endoglucanase activity (including a variant) comprises a CBM domain (*e.g.*, functional CBM domain). In some aspects, the polypeptide having GH61/endoglucanase activity (including a variant of GH61 endoglucanase) comprises a catalytic domain (*e.g.*, functional catalytic domain).

[0025] Also provided herein are variants of EG IV polypeptides. For example, such variants can have at least about 60% (*e.g.*, at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) sequence identity to any one of SEQ ID NOs: 1-29 and 148, or to a mature polypeptide thereof. For example, provided herein are variants of *T. reesei* Eg4 polypeptide. Such variants may have at least about 60% (*e.g.*, at least about 60%, 65%, 70%, 75%, 80%, 85%, 88%, 90%, 92.5%, 95%, 96%, 97%, 98%, or 99%) sequence identity to residues 22 to 344 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof is isolated. In some aspects, the polypeptide or a variant thereof has endoglucanase activity. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to at least about 5 residues (*e.g.*, at least about any of 6, 7, 8, 9, 10, 11, or 12) of H22, D61, G63, C77, H107, R177, E179, H184, Q193, C198, Y195, and Y232 of SEQ ID NO:27, or any corresponding conserved residues in any of the other polypeptides. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to H22, D61, G63, C77, H107, R177, E179, H184, Q193, C198, Y195, and Y232 of SEQ ID NO:27. The polypeptide or a variant thereof may comprise residues corresponding to at least 5 residues (*e.g.*, at least about any of 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, or 19) of G313, Q314, C315, G316, G317, S321, G322, P323, T324, C326, A327, T331, C332, N336, Y338, Y339, Q341, C342, and L343 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to G313, Q314, C315, G316, G317, S321, G322, P323, T324, C326, A327, T331, C332, N336, Y338, Y339, Q341, C342, and L343 of SEQ ID NO:27. The polypeptide or a variant thereof may comprise a CBM domain (*e.g.*, a functional CBM domain). In some aspects, the polypeptide or a variant thereof comprises a catalytic domain (*e.g.*, a functional catalytic domain).

[0026] Also provided herein are nucleic acids or polynucleotides encoding any one of the polypeptides herein. For example, the disclosure provides polynucleotide encoding a

polypeptide having at least about 60% (*e.g.*, at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) sequence identity to any one of SEQ ID NOs: 1-29 and 148. For example, the disclosure provides herein isolated nucleic acids having at least about 60% (*e.g.*, at least about 60%, 65%, 70%, 75%, 80%, 85%, 88%, 90%, 92.5%, 95%, 96%, 97%, 98%, or 99%) identity to SEQ ID NO:30. Also provided are expression cassettes, vectors, and cells comprising the nucleic acids described above.

[0027] Also provided herein are enzyme compositions (*e.g.*, non-naturally occurring compositions) comprising a polypeptide having GH61/endoglucanase activity. In some aspects, the composition comprises a whole cellulase comprising the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof). The polypeptide having GH61/ endoglucanase activity is, *e.g.*, *T. reesei* endoglucanase IV (“*T. reesei* Eg4”) or a variant thereof. A variant of *T. reesei* Eg4 can be any of the variants provided herein.

[0028] In some aspects, the enzyme composition is a cellulase composition. The enzyme composition may further comprise one or more hemicellulases, and thus can also be a hemicellulase composition. In some aspects, the enzyme composition comprises at least one (*e.g.*, at least 2, 3, 4, 5, 6, 7, or 8) cellulase polypeptide(s). In some aspects, the at least one cellulase polypeptide is a polypeptide having endoglucanase activity, a polypeptide having cellobiohydrolase activity, or a polypeptide having β -glucosidase activity. In some aspects, the composition further comprises at least one (*e.g.*, at least 2, 3, 4, 5, 6, 7, or 8)

hemicellulase polypeptide(s). In some aspects, the at least one hemicellulase polypeptide is a polypeptide having xylanase activity, a polypeptide having β -xylosidase activity, or a polypeptide having L- α -arabinofuranosidase activity, or a polypeptide having combined xylanase/ β -xylosidase activity, combined β -xylosidase/L- α -arabinofuranosidase activity, or combined xylanase/ L- α -arabinofuranosidase activity activity. In some aspects, the composition comprises at least one (*e.g.*, at least 2, 3, 4, 5, 6, 7, or 8) cellulase polypeptide(s) and at least one (*e.g.*, at least 2, 3, 4, 5, 6, 7, or 8) hemicellulase polypeptide(s).

[0029] In some aspects, the enzyme composition comprises a polypeptide having GH61/ endoglucanase activity and further comprises at least 1 (*e.g.*, at least 2, 3, 4, or 5) polypeptide having endoglucanase activity, at least 1 (*e.g.*, at least 2, 3, 4, or 5) polypeptide having cellobiohydrolase activity, at least 1 (*e.g.*, at least 2, 3, 4, or 5) polypeptide having β -glucosidase activity, at least 1 (*e.g.*, at least 2, 3, 4, or 5) polypeptide having xylanase activity, at least 1 (*e.g.*, at least 2, 3, 4, or 5) polypeptide having β -xylosidase activity, and/or

at least 1 (*e.g.*, at least 2, 3, 4, or 5) polypeptide having L- α -arabinofuranosidase activity.

[0030] In some aspects, the composition comprises a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one polypeptide having xylanase activity (*e.g.*, *T. reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, AfuXyn5, or a variant thereof). In some aspects, the composition further comprises at least one polypeptide having β -glucosidase activity (*e.g.*, Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B, or a variant thereof). In some aspects, the composition further comprises at least one polypeptide having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof). In some aspects, the composition further comprises at least one polypeptide having endoglucanase activity other than the GH61 enzyme (*e.g.*, *T. reesei* EG1, *T. reesei* EG2, or a variant thereof).

[0031] The composition may comprise a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least 1 polypeptide having β -glucosidase activity (*e.g.*, Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B or a variant thereof). The composition may comprise a polypeptide having GH61/endoglucanase activity and at least 1 polypeptide having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B or a variant thereof). The composition may comprise a polypeptide having GH61/endoglucanase activity, and at least 1 polypeptide having endoglucanase activity (*e.g.*, *T. reesei* EG1, *T. reesei* EG2 or a variant thereof). The composition may comprise a polypeptide having GH61/endoglucanase activity and at least 1 polypeptide having β -xylosidase activity (*e.g.*, Fv3A, Fv43A, Pf43A, Fv43D, Fv39A, Fv43E, Fo43A, Fv43B, Pa51A, Gz43A, *T. reesei* Bxl1 or a variant thereof). The composition may comprise a polypeptide having GH61/endoglucanase activity and at least 1 polypeptide having L- α -arabinofuranosidase activity (*e.g.*, Af43A, Fv43B, Pf51A, Pa51A, Fv51A or a variant thereof).

[0032] Any one of the compositions described herein may comprise a whole cellulase. For example, a composition is provided comprising a whole cellulase comprising a polypeptide having GH61/endoglucanase activity. Alternatively, a composition is provided comprising a whole cellulase plus a polypeptide having GH61/endoglucanase activity. In some aspects, a composition comprising a polypeptide having GH61/endoglucanase activity, and a

polypeptide having endoglucanase activity other than the polypeptide having GH61/ endoglucanase activity, a polypeptide having cellobiohydrolase activity, and a polypeptide having β -glucosidase activity is provided. The composition further comprises one or more hemicellulase polypeptides. For example, the composition may comprise one or more

5 polypeptides having xylanase activity, one or more polypeptides having β -xylosidase activity, and/or one or more polypeptides having L- α -arabinofuranosidase activity. A composition may comprise a polypeptide having GH61/endoglucanase activity, at least one polypeptide having xylanase activity (e.g., *T. reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, AfuXyn5, or a variant thereof), and a whole cellulase. In some aspects, a composition comprising a

10 polypeptide having GH61/ endoglucanase activity, at least one polypeptide having xylanase activity (e.g., *T. reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, AfuXyn5, or a variant thereof), and at least one other polypeptide having hemicellulase activity is provided.

[0033] In some aspects, the whole cellulase comprises at least one polypeptide having endoglucanase activity (e.g., *T. reesei* EG1, *T. reesei* EG2, or a variant thereof) that is not the

15 polypeptide having GH61/endoglucanase activity. The whole cellulase can comprise at least one polypeptide having cellobiohydrolase activity (e.g., *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof). The whole cellulase can comprise at least one polypeptide having β -glucosidase activity (e.g., Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B,

20 Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B, or a variant thereof).

[0034] In some aspects, in any one of the compositions described herein, the at least one polypeptide having endoglucanase activity but is not the one having GH61/endoglucanase activity is, e.g., *T. reesei* EG1 (or a variant thereof) and/or *T. reesei* EG2 (or a variant thereof). In some aspects, the at least one polypeptide having cellobiohydrolase activity is,

25 e.g., *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof. In some aspects, the at least one polypeptide having β -glucosidase activity is, e.g., Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, and/or Tn3B, or variants thereof. In some aspects, the at least one polypeptide having xylanase activity is, e.g., *T.*

30 *reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, and/or AfuXyn5, or variants thereof. In some aspects, the at least one polypeptide having β -xylosidase activity is, e.g., a Group 1 β -xylosidase or a Group 2 β -xylosidase, wherein the Group 1 β -xylosidase may be Fv3A,

Fv43A polypeptide, or a variant thereof, and the Group 2 β -xylosidase may be Pf43A, Fv43D, Fv39A, Fv43E, Fo43A, Fv43B, Pa51A, Gz43A, *T. reesei* Bxl1 polypeptide, or a variant thereof. In some aspects, the at least one polypeptide having β -xylosidase activity is, e.g., Fv3A (or a variant thereof) and/or Fv43D (or a variant thereof). In some aspects, the at least one polypeptide having L- α -arabinofuranosidase activity may be Af43A, Fv43B, Pf51A, Pa51A, and/or Fv51A, or variants thereof.

[0035] In some aspects, a composition comprising an isolated polypeptide having GH61/endoglucanase activity (e.g., *T. reesei* Eg4 or a variant thereof) is provided. In some aspects, the polypeptide having GH61/endoglucanase activity (e.g., *T. reesei* Eg4 or a variant thereof) is expressed by a host cell, wherein the nucleic acid encoding the polypeptide having GH61/endoglucanase activity has been engineered into the host cell. For example, the polypeptide having GH61/endoglucanase activity is expressed by a host cell, and the nucleic acid encoding that polypeptide is heterologous to the host cell.

[0036] In some aspects, a composition is provided comprising a polypeptide having GH61/endoglucanase activity (e.g., *T. reesei* Eg4 or a variant thereof), and further comprising one or more cellulase polypeptides and/or one or more hemicellulase polypeptides, wherein the cellulase polypeptide and/or the hemicellulase polypeptide is expressed by a host cell, and the cellulase polypeptide and/or hemicellulase polypeptide is heterologous to the host cell. In some aspects, a composition comprising a polypeptide having GH61/endoglucanase activity and further comprising at least one cellulase polypeptide and/or at least one hemicellulase polypeptide is provided, and the cellulase polypeptide and/or the hemicellulase polypeptide is expressed by a host cell, and the cellulase polypeptide and/or hemicellulase polypeptide is endogenous to the host cell. In some aspects, the cellulase polypeptide comprises a polypeptide having endoglucanase activity (e.g., *T. reesei* EG1, *T. reesei* EG2, or a variant thereof) that is different from the polypeptide having GH61/endoglucanase activity, a polypeptide having cellobiohydrolase activity (e.g., *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof), or a polypeptide having β -glucosidase activity (e.g., Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B, or a variant thereof). In some aspects, the hemicellulase polypeptide comprises a polypeptide having xylanase activity (e.g., *T. reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, AfuXyn5, or a variant thereof), a polypeptide having β -xylosidase activity (e.g., Fv3A, Fv43A, Pf43A,

Fv43D, Fv39A, Fv43E, Fo43A, Fv43B, Pa51A, Gz43A, *T. reesei* Bx11, or a variant thereof), or a polypeptide having L- α -arabinofuranosidase activity (*e.g.*, Af43A, Fv43B, Pf51A, Pa51A, Fv51A, or a variant thereof).

[0037] In some aspects, the composition is prepared from a fermentation broth. In some

5 aspects, the composition is prepared from the fermentation broth of an integrated strain (*e.g.*, H3A/Eg4, #27, as described herein in the Examples), wherein the GH61 endoglucanase gene is integrated into the genetic materials of the host strain. In some aspects, the composition is prepared from the fermentation broth of a strain, wherein a nucleic acid encoding a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is
10 heterologous to the host cell, wherein the GH61 endoglucanase has been, *e.g.*, integrated into the strain, or expressed by a vector introduced into the host strain.

[0038] Any one of the compositions or methods provided herein comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) may be a whole cellulase. The composition may be a fermentation broth subject to minimum post-production
15 processing (*e.g.*, purification, filtration, a cell kill step, and/or ultrafiltration, etc), and is used as a whole broth formulation.

[0039] In some aspects, a composition (*e.g.*, a non-naturally occurring composition) is provided comprising *T. reesei* Eg4, *T. reesei* Bgl1, *T. reesei* xyn3, Fv3A, Fv43D, and Fv51A, or respective variants thereof. The composition may be a whole cellulase. The composition
20 may be a fermentation broth subject to minimum post-production processing (*e.g.*, filtration, purification, ultrafiltration, a cell-kill step, etc), and is thus used as a whole broth formulation. In some aspects, the composition comprises an isolated *T. reesei* Eg4 or a variant thereof. In some aspects, the composition comprises at least one of an isolated *T. reesei* Bgl1, an isolated *T. reesei* xyn3, an isolated Fv3A, an isolated Fv43D, and an isolated Fv51A. For example,
25 any of the above-mentioned polypeptides can be introduced into the composition by simple addition or mixing of purified or isolated polypeptides. Alternatively, the polypeptides herein can be expressed by the host strain using suitable recombinant techniques, and certain of the above-mentioned polypeptides may be overexpressed or underexpressed, as compared to their naturally-occurring levels in the host cell. In some aspects, genes encoding any one of
30 the above-mentioned polypeptides can be integrated into the host strain. In some aspects, the composition of the present disclosure is prepared from a fermentation broth of the host strain. In some aspects, the composition is from the fermentation broth of an integrated strain (*e.g.*,

H3A/Eg4, #27, as described herein in the Examples). In some embodiments, the fermentation broth is subject to minimum post-production processing, and is used as a whole broth formulation. In some aspects, the nucleic acid encoding the GH61 endoglucanase is heterologous to the host cell. In some aspects, at least one of the nucleic acids encoding *T. reesei* Bgl1, *T. reesei* xyn3, Fv3A, Fv43D, or Fv51A is heterologous to the host cell expressing the GH61 endoglucanase of the invention. In some aspects, at least one nucleic acid encoding *T. reesei* Bgl1, *T. reesei* xyn3, Fv3A, Fv43D, or Fv51A is endogenous to the host cell expressing the GH61 endoglucanase.

[0040] The polypeptide having GH61/endoglucanase activity (e.g., *T. reesei* Eg4 or a variant thereof) may be present in an enzyme composition or in a biomass saccharification mixture in an amount sufficient to increase the yield of fermentable sugar(s) from hydrolysis of a biomass material (e.g., by at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%) as compared to the yield achieved by a control enzyme composition or a control biomass saccharification mixture that is comparable in terms of the types and concentrations of enzymatic or other components therein, but without the polypeptide(s) having GH61/endoglucanase activity. The polypeptide having GH61/endoglucanase activity may be present in the enzyme composition or in a biomass saccharification mixture in an amount sufficient to reduce the viscosity of the biomass saccharification mixture during hydrolysis of the biomass material therein (e.g., by at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%) as compared to the viscosity of a control mixture that is comparable in terms of the types and concentrations of enzymatic or other components therein, but without the polypeptide having GH61/endoglucanase activity. In some aspects, the enzyme composition or the biomass saccharification mixture comprises at least 1 polypeptide having endoglucanase activity, at least 1 polypeptide having cellobiohydrolase activity, at least 1 polypeptide having β -glucosidase activity, in total amounts that are sufficient to cause hydrolysis of the biomass material to which the polypeptides come into contact. The enzyme composition or the biomass saccharification mixture may further comprise at least 1 polypeptide having xylanase activity, at least 1 polypeptide having β -xylosidase activity, at least 1 polypeptide having L- α -arabinofuranosidase activity, and/or a whole cellulase, or a mixture thereof, in total amounts that are sufficient to cause hydrolysis of the biomass material to which the polypeptides come into contact.

[0041] In some aspects, the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is present in an amount that is about 0.1 wt.% to about 50 wt.% (*e.g.*, about 0.5 wt.% to about 30 wt.%, about 1 wt.% to about 20 wt.%, about 5 wt.% to about 20 wt.%, about 7 wt.% to about 20 wt.%, or about 8 to about 15 wt.%) of the total weight of proteins in the enzyme composition or in the biomass saccharification mixture. For example the polypeptide having GH61/endoglucanase activity is present in an amount that is about 8 wt.%, about 10 wt.%, or about 12 wt.% of the total weight of proteins in the enzyme composition or in the biomass saccharification mixture. The enzyme composition or the biomass saccharification mixture may comprise more than one polypeptides having GH61/endoglucanase activity. For example, the enzyme composition or biomass saccharification mixture can comprise a *T. reesei* Eg4 or a variant thereof, as well as a *T. reesei* Eg7 (or a variant thereof), wherein the total amount of polypeptides having GH61/endoglucanase (Eg4 + Eg7) activity is about 0.1 wt.% to about 50 wt.% (*e.g.*, about 0.5 wt.% to about 30 wt.%, about 2 wt.% to about 20 wt.%, about 5 wt.% to about 20 wt.%, about 7 wt.% to about 20 wt.%, or about 8 wt.% to about 15 wt.%) of the total weight of proteins in the enzyme composition or in the biomass saccharification mixture. The polypeptide(s) having GH61/endoglucanase activity may be expressed from polynucleotides that are heterologous or endogenous to the host cell. Alternatively the polypeptide having GH61/endoglucanase activity can be introduced into the enzyme composition or the biomass saccharification mixture in an isolated or purified form.

[0042] In some aspects, a polypeptide having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof) is present in an amount that is about 0.1 wt.% to about 80 wt.% (*e.g.*, about 5 wt.% to about 70 wt.%, about 10 wt.% to about 60 wt.%, about 20 wt.% to about 50 wt.%, or about 25 wt.% to about 50 wt.%) of the total weight of proteins in the enzyme composition or the biomass saccharification mixture. The enzyme composition or biomass saccharification mixture may comprise more than one polypeptide having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof), wherein the total amount of polypeptides having cellobiohydrolase activity is about 0.1 wt.% to about 80 wt.% (*e.g.*, about 5 wt.% to about 70 wt.%, about 10 wt.% to about 60 wt.%, about 20 wt.% to about 50 wt.%, or about 25 wt.% to about 50 wt.%)

of the total weight of proteins in the enzyme composition or the biomass saccharification mixture. The polypeptide having cellobiohydrolase activity is, in some aspects, expressed from a nucleic acid heterologous or endogenous to the host cell. In some aspects, the polypeptide having cellobiohydrolase activity can be introduced into the enzyme composition or biomass saccharification mixture in an isolated or purified form.

[0043] The enzyme composition or the biomass saccharification mixture may comprise one or more polypeptides having β -glucosidase activity (*e.g.*, Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B or a variant thereof), wherein the total amount of polypeptides having β -glucosidase activity is about 0.1 wt.% to about 50 wt.% (*e.g.*, about 1 wt.% to about 30 wt.%, about 2 wt.% to about 20 wt.%, about 5 wt.% to about 20 wt.%, or about 8 wt.% to about 15 wt.%) of the total weight of proteins in the enzyme composition or biomass saccharification mixture. The polypeptide having β -glucosidase activity may be expressed from a nucleic acid heterologous or endogenous to the host cell. The polypeptide having β -glucosidase activity may alternatively be introduced into the enzyme composition or biomass saccharification mixture in an isolated or purified form.

[0044] In some aspects, the enzyme composition or biomass saccharification mixture can comprise one or more the polypeptides having xylanase activity (*e.g.*, *T. reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, AfuXyn5, or a variant thereof), wherein the total amount of polypeptides having xylanase activity is about 0.1 wt.% to about 50 wt.% (*e.g.*, about 1 wt.% to about 40 wt.%, about 4 wt.% to about 30 wt.%, about 5 wt.% to about 20 wt.%, or about 8 wt.% to about 15 wt.%) of the total weight of proteins in the enzyme composition or the biomass saccharification mixture. The polypeptide having xylanase activity can be expressed from a nucleic acid heterologous or endogenous to the host cell. In some aspects, the polypeptide having xylanase activity can be introduced or mixed into the enzyme composition or the biomass saccharification mixture in an isolated or purified form.

[0045] The enzyme composition or biomass saccharification mixture may comprise one or more polypeptides having L- α -arabinofuranosidase activity (*e.g.*, Af43A, Fv43B, Pf51A, Pa51A, Fv51A, or a variant thereof), wherein the total amount of polypeptides having L- α -arabinofuranosidase activity is about 0.1 wt.% to about 50 wt.% (*e.g.*, about 1 wt.% to about 40 wt.%, about 2 wt.% to about 30 wt.%, about 4 wt.% to about 20 wt.%, or about 5 wt.% to about 15 wt.%) of the total weight of proteins in the enzyme composition or the biomass saccharification mixture. The polypeptide having L- α -arabinofuranosidase activity may be

expressed from a nucleic acid heterologous or endogenous to the host cell. In some aspects, the polypeptide having L- α -arabinofuranosidase activity can be introduced or mixed into the enzyme composition or the biomass saccharification mixture in an isolated or purified form.

[0046] The enzyme composition or the biomass saccharification mixture may comprise one

5 or more polypeptides having β -xylosidase activity (e.g., Fv3A, Fv43A, Pf43A, Fv43D, Fv39A, Fv43E, Fo43A, Fv43B, Pa51A, Gz43A, *T. reesei* Bx11 or a variant thereof), wherein the total amount of the polypeptides having β -xylosidase activity is about 0.1 wt.% to about 50 wt.% (e.g., about 1 wt.% to about 40 wt.%, about 4 wt.% to about 35 wt.%, about 5 wt.% to about 25 wt.%, or about 5 wt.% to about 20 wt.%) of the total weight of proteins in the enzyme

10 composition or the biomass saccharification mixture. The polypeptide having β -xylosidase activity may be expressed from a nucleic acid heterologous or endogenous to the host cell.

The polypeptide having β -xylosidase activity may alternatively be introduced into the enzyme composition or the biomass saccharification mixture in an isolated or purified form.

[0047] In some aspects, the enzyme composition provided herein may be a whole cellulase.

15 The whole cellulase may comprise one or more polypeptides having endoglucanase activity (such as, e.g., *T. reesei* Eg4, Eg1, Eg2, Eg7, or a variant thereof) expressed from a nucleic acid heterologous or endogenous to the host cell. The whole cellulase may also comprise one or more polypeptides having cellobiohydrolase activity (e.g., *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S.*

20 *thermophile* 6A, 6B, or a variant thereof) expressed from a nucleic acid heterologous or endogenous to the host cell. The whole cellulase may further comprise one or more polypeptide having β -glucosidase activity (e.g., Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B, or a variant thereof) expressed from a nucleic acid heterologous or endogenous to the host cell. The whole cellulase may be used

25 in the form of a fermentation broth of the host cell. The broth can be subject to minimum post-production processing, including, e.g., filtration, purification, ultrafiltration, a cell-kill step, etc, and thus the broth may be used for biomass hydrolysis in a whole broth formulation.

[0048] In some aspects, the enzyme composition provided herein is capable of converting a biomass material into fermentable sugar(s) (e.g., glucose, xylose, arabinose, and/or

30 cellobiose). In some aspects, the enzyme composition is capable of achieving at least about 0.1 (e.g., 0.1 to 0.4) fraction product as determined by the calcofluor assay described herein.

[0049] In some aspects, the enzyme composition can be a cellulase composition or a

hemicellulase composition. The enzyme composition may comprise the polypeptide having GH61/endoglucanase activity and further may comprise one or more cellulase polypeptides and/or one or more hemicellulase polypeptides, wherein the one or more polypeptides having GH61/endoglucanase activity and the one or more cellulase polypeptides, and/or the one or more hemicellulase polypeptides are blended into a mixture before the mixture is used to contact and hydrolyze a biomass substrate in a biomass saccharification mixture.

[0050] In some aspects, the one or more polypeptides having GH61/endoglucanase activity, one or more cellulase polypeptides, and one or more hemicellulase polypeptide, are added to a biomass material, at different times. For example, a polypeptide having GH61/endoglucanase activity is added to a biomass material before, or after, a cellulase polypeptide and/or a hemicellulase polypeptide is added to the same biomass material.

[0051] In some aspects, a composition of the invention comprises at least one polypeptide having GH61/endoglucanase activity and a biomass material in, e.g., a mixture. For example, the composition may be a hydrolysis mixture, a fermentation broth/mixture, or a biomass saccharification mixture. The mixture may comprise one or more fermentable sugar(s).

[0052] Also provided herein are methods of hydrolyzing a biomass material comprising contacting the biomass material with an enzyme composition (e.g., a non-naturally occurring composition) comprising a polypeptide having GH61/endoglucanase activity, in an amount sufficient to hydrolyze the biomass material in the resulting biomass saccharification mixture.

[0053] Also provided herein are methods of reducing the viscosity of a biomass mixture, and/or a biomass saccharification mixture comprising contacting the mixture with an enzyme composition (e.g., a non-naturally occurring composition) comprising a polypeptide having GH61/endoglucanase activity, which is present in the composition in an amount sufficient to reduce the viscosity of the mixture. In some aspects, the biomass mixture or the biomass saccharification mixture comprises a biomass material, optionally also fermentable sugar(s), a whole cellulase and/or a composition comprising a polypeptide having cellulase activity and/or a polypeptide having hemicellulase activity. The viscosity of the mixture may be reduced by at least about 5%, (e.g., at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%) as compared to the viscosity of a control mixture comprising the same components at the same concentrations except that the polypeptide having GH61/endoglucanase activity is absent from the mixture. The biomass material may comprise hemicellulose, cellulose, or a mixture thereof. The biomass material may

comprises glucan, xylan and/or lignin, or a mixture thereof.

[0054] In some aspects, the biomass material can suitably be treated or pre-treated with an acid or a base. In some aspects, the base is ammonia. The method of the invention may further comprise adjusting the pH of the biomass mixture to a pH of about 4.0 to about 6.5 (e.g., pH of about 4.5 to about 5.5). In some aspects, the method is performed at a pH of about 4.0 to about 6.5 (e.g., pH of about 4.5 to about 5.5). In some aspects, the method is performed for about 2 h to about 7 d (e.g., about 4 h to about 6 d, about 8 h to about 5 d, or about 8 h to about 3 d). This pH adjustment can suitably be made before putting the biomass mixture in contact with the polypeptides or the enzyme compositions.

[0055] In some aspects, the biomass material is present in a saccharification mixture in a high solids level, e.g., the biomass material in its solid state constitutes at least about 5 wt.% to about 60 wt.% (e.g., about 10 wt.% to about 50 wt.%, about 15 wt.% to about 40 wt.%, about 15 wt.% to about 30 wt.%, or about 20 wt.% to about 30 wt.%) of the total weight of enzymes plus biomass materials in the saccharification mixture. By the weight of the biomass material in its solid state, it is meant the weight of the biomass material in its dry state, its dry solid state, its natural state, or its unprocessed state, or before the biomass is contacted with the polypeptides in the enzyme composition. Preferably the biomass material in its solid state constitutes at least about 15 wt.%, and even more preferably at least about 20 wt.% or 25 wt.% of the total weight of enzymes plus biomass materials in the saccharification mixture.

[0056] In some aspects, the method comprises producing fermentable sugar(s). The amount of fermentable sugar(s) may be produced at an increased level using the method of the invention. For example, the amount of the fermentable sugar(s) produced using the methods or the compositions herein is increased by at least about 5% (e.g., at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%) as compared to the amount of the fermentable sugar(s) produced when the same biomass material is hydrolyzed by an enzyme composition comprising the same polypeptide components at the same concentrations, except that polypeptide having GH61/endoglucanase activity is absent.

[0057] In some aspects, the amount of the enzyme composition comprising a polypeptide having GH61/endoglucanase activity is sufficient to increase the yield of fermentable sugar(s) by at least about 5%, (e.g., at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%), as compared to the yield of fermentable sugar(s) from the same biomass material by an enzyme composition having the same components at

the same concentrations, except that the polypeptide having GH61/endoglucanase activity is absent. In some aspects, the amount of the polypeptide having GH61/endoglucanase activity in the biomass saccharification mixture is sufficient to reduce the viscosity of the mixture by at least about 5% (*e.g.*, at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%,
5 60%, 70%, 80%, or 90%) as compared to the viscosity of a control biomass saccharification mixture comprising the same biomass and the same panel of polypeptides at the same concentrations, except that the polypeptide having GH61/endoglucanase activity is absent.

[0058] In some aspects, the amount of the composition comprising a polypeptide having GH61/endoglucanase activity used in a saccharification or hydrolysis process is about 0.1 mg
10 to about 50 mg protein (*e.g.*, about 0.2 mg to about 40 mg protein, about 0.5 mg to about 30 mg protein, about 1 mg to about 20 mg protein, or about 5 mg to about 15 mg protein) per gram of cellulose, hemicellulose, or a mixture of cellulose and hemicelluloses in the biomass material. The protein amount described herein refers to the weight of total protein in the enzyme composition or the biomass saccharification mixture. The proteins include a
15 polypeptide having GH61/endoglucanase activity and may include other enzymes such as cellulase polypeptide(s) and/or hemicellulase polypeptide(s). In some aspects, the amount of the polypeptide having GH61/endoglucanase activity used in the hydrolysis or saccharification process is about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 1 mg to about 5 mg) protein per gram of cellulose,
20 hemicellulose, or cellulose and hemicelluloses contained in the biomass material.

[0059] The enzyme composition or biomass saccharification mixture comprising a polypeptide having GH61/endoglucanase activity and at least 1 polypeptide having endoglucanase activity (*e.g.*, *T. reesei* Eg1, *T. reesei* Eg2, and/or a variant thereof) in the hydrolysis or saccharification process may contain about 0.2 mg to about 30 mg (*e.g.*, about
25 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 1 mg to about 5 mg) protein per gram of cellulose, hemicellulose, or cellulose and hemicellulose in the biomass material.

[0060] The enzyme composition or biomass saccharification mixture comprising a polypeptide having GH61/endoglucanase activity and at least 1 polypeptide having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B,
30 *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof) in the hydrolysis or saccharification process may contain about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 1 mg to about

5 mg) protein per gram of cellulose, hemicellulose, or cellulose and hemicellulose in the biomass material.

[0061] In some aspects, the enzyme composition or biomass saccharification mixture comprising a polypeptide having GH61/endoglucanase activity and at least 1 polypeptide having β -glucosidase activity (*e.g.*, Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B, or a variant thereof) in the hydrolysis or saccharification process may contain about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 0.5 mg to about 5 mg) protein per gram of cellulose, hemicellulose, or cellulose and hemicellulose in the biomass material.

[0062] The enzyme composition or biomass saccharification mixture comprising a polypeptide having GH61/endoglucanase activity and at least 1 polypeptide having xylanase activity (*e.g.*, *T. reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, AfuXyn5 or a variant thereof) in the hydrolysis or saccharification process may contain about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, about 0.5 mg to about 5 mg) protein per gram of cellulose, hemicellulose, or cellulose and hemicellulose in the biomass material.

[0063] The enzyme composition or the biomass saccharification mixture comprising a polypeptide having GH61/endoglucanase activity and at least 1 polypeptide having β -xylosidase activity (*e.g.*, Fv3A, Fv43A, Pf43A, Fv43D, Fv39A, Fv43E, Fo43A, Fv43B, Pa51A, Gz43A, *T. reesei* Bx11, and/or a variant thereof) used in the hydrolysis or saccharification process may contain about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 0.5 mg to about 5 mg) protein per gram of cellulose, hemicellulose, or cellulose and hemicellulose in the biomass material.

[0064] The enzyme composition or the biomass saccharification mixture comprising a polypeptide having GH61/endoglucanase activity and at least 1 polypeptide having L- α -arabinofuranosidase activity (*e.g.*, Af43A, Fv43B, Pf51A, Pa51A, Fv51A, and/or a variant thereof) used in the hydrolysis or saccharification process may contain about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 0.5 mg to about 5 mg) protein per gram of cellulose, hemicellulose, or cellulose and hemicellulose in the biomass material.

[0065] In some aspects, the method of the invention is performed at a temperature of about 30°C to about 65°C (*e.g.*, about 35°C to about 60°C, about 40°C to about 60°C, or about 45°C to about 55°C).

[0066] The method of the invention may further comprise the step of contacting the biomass material with an enzyme composition comprising a whole cellulase. In some aspects, the step of further contacting the biomass material with a composition comprising a whole cellulase is performed before, after, or concurrently with contacting the biomass material with an enzyme composition comprising a polypeptide having GH61/endoglucanase activity.

[0067] In some aspects, the method of the invention further comprises the step contacting the biomass material with an enzyme composition comprising a polypeptide having cellulase activity and/or a polypeptide having hemicellulase activity. The step of contacting the biomass material with a composition comprising a polypeptide having cellulase activity and/or a polypeptide having hemicellulase activity may be performed before, after, or concurrently with contacting the biomass material with an enzyme composition comprising a polypeptide having GH61/endoglucanase activity.

[0068] In some aspect, the composition comprises the polypeptide having GH61/endoglucanase activity and further comprises at least 1 cellulase polypeptide and/or at least one hemicellulase polypeptide, wherein the polypeptide having GH61/endoglucanase activity and at least one cellulase polypeptide and/or at least 1 hemicellulase polypeptide are blended into a mixture before the mixture is used to contact the biomass material.

[0069] In some aspects, the composition comprises the polypeptide having GH61/endoglucanase activity and further comprises 1 or more cellulase polypeptides and/or 1 or more hemicellulase polypeptides, wherein the polypeptide having GH61/endoglucanase activity and 1 or more cellulase polypeptides and/or 1 or more hemicellulase polypeptides are added to the biomass material at different times. For example, the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is added before or after the 1 or more cellulase polypeptides and/or the 1 or more hemicellulase polypeptides are added.

[0070] In some aspects, methods of applying the invention in both an industrial setting and/or a commercial setting are contemplated. Accordingly a method or a method of manufacturing, marketing, or otherwise commercializing the instant compositions comprising suitable GH61 endoglucanases is within the purview of the disclosure. The method includes, for example, the application of the compositions or the GH61 endoglucanase polypeptides or variants thereof in a merchant enzyme supply model, wherein the enzymes and variants, as well as the compositions of the invention are supplied or sold to cellulosic sugar producers, certain ethanol (bioethanol) refineries or other bio-chemical or bio-material manufacturers.

The method can also be, in some aspects, the application of the compositions or the GH61 endoglucanase polypeptides or variants thereof in an on-site bio-refinery model, wherein the polypeptides or variants, or the non-naturally occurring cellulase and hemicellulase compositions of the invention are produced in an enzyme production system that is built by the enzyme manufacturer at a site that is located at or in the vicinity of the cellulosic sugar plant, bioethanol refineries or the bio-chemical/biomaterial manufacturers. In some aspects, suitable biomass substrates, preferably subject to appropriate pretreatments as described herein, can be hydrolyzed using the saccharification methods and the enzymes and/or enzyme compositions herein at or near the bioethanol refineries or the bio-chemical/biomaterial manufacturing facilities. The resulting fermentable sugars can then be subject to fermentation at the same facilities or at facilities in the vicinity.

[0071] It is to be understood that one, some, or all of the properties of the embodiments described herein may be combined to form other embodiments of the present invention. These and other aspects of the invention will become apparent to one of skill in the art.

BRIEF DESCRIPTION OF THE FIGURES

[0072] The skilled artisan will understand that the drawings are for illustration purposes only and are not intended to limit the scope of the present teachings in anyway.

[0073] **FIG. 1:** depicts certain amino acid sequences of various polypeptides having GH61/endoglucanase activity.

[0074] **FIG. 2:** depicts percent identity and divergence using ClustalV (PAM250) comparing a number of amino acid sequences of various polypeptides having GH61/endoglucanase activity, such as those presented in FIG. 1 (SEQ ID NOs: 1-28).

[0075] **FIG. 3:** depicts the alignment of various polypeptides having GH61/endoglucanase activity such as those presented in FIG. 1 (SEQ ID NOs: 1-28).

FIGs. 4A-4B: **FIG. 4A** depicts nucleotide sequence of *T. reesei* Eg4 (SEQ ID NO:30). **FIG. 4B** depicts amino acid sequence of *T. reesei* Eg4 (SEQ ID NO:27). The predicted signal sequence is underlined, the predicted conserved domains are in bold, and the predicted linker is in italic.

[0077] **FIG. 5:** depicts an amino acid sequence alignment of *T. reesei* Eg4 (TrEG4) (SEQ ID NO:27) with *T. reesei* Eg7 (TrEG7, or TrEGb) (SEQ ID NO:26) and TtEG (SEQ ID NO:29).

[0078] **FIGs. 6A-6B:** **FIG. 6A** provides conserved residues of *T. reesei* Eg4 (TrEg4), inferred from sequence alignment and the known structures of TrEG7 (crystal structure at Protein Data Bank Accession: pdb:2vtc) and TtEG (crystal structure at Protein Data Bank Accession: pdb:3EII). **FIG. 6B** provides conserved CBM domain residues inferred from sequence alignment with known sequences of Tr6A, and Tr7A.

[0079] **FIG. 7** lists a number of amino acid sequence motifs of GH61 endoglucanases. Each of the “a”s in the sequence motifs represents an amino acid that may be any one of alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, or valine.

[0080] **FIGs. 8A-8I:** **FIG. 8A** depicts pENTR-TOPO-Bgl1-943/942 plasmid. **FIG. 8B** depicts pTrex3g 943/942 expression vector. **FIG. 8C** depicts pENTR/*T. reesei* Xyn3 plasmid. **FIG. 8D** depicts pTrex3g/*T. reesei* Xyn3 expression vector. **FIG. 8E** depicts pENTR-Fv3A plasmid. **FIG. 8F** depicts the pTrex6g plasmid. **FIG. 8G** depicts pTrex6g/Fv3A expression vector. **FIG. 8H** depicts TOPO Blunt/Pgl1-Fv43D plasmid. **FIG. 8I** depicts TOPO Blunt/Pgl1-Fv51A plasmid.

[0081] **FIG. 9:** provides the enzyme composition of *T. reesei* integrated strain H3A.

[0082] **FIG. 10:** lists the enzymes (purified or unpurified) that were individually added to each of the samples in Example 2, and the stock protein concentrations of these enzymes.

[0083] **FIG. 11A-11D:** **FIG. 11A** depicts glucose release following saccharification of dilute ammonia pretreated corncob by adding enzyme compositions comprising various purified or non-purified enzymes of **FIG. 10**, which were added to *T. reesei* integrated strain H3A, in accordance with Example 2. **FIG. 11B** depicts cellobiose release following saccharification of dilute ammonia pretreated corncob by adding enzyme compositions comprising various purified or non-purified enzymes of **FIG. 10**, which were added to *T. reesei* integrated strain H3A, in accordance with Example 2; **FIG. 11C** depicts xylobiose release following saccharification of dilute ammonia pretreated corncob by adding enzyme compositions comprising various purified or non-purified enzymes of **FIG. 10**, which were added to *T. reesei* integrated strain H3A, in accordance with Example 2; **FIG. 11D** depicts xylose release following saccharification of dilute ammonia pretreated corncob by adding enzyme compositions comprising various purified or non-purified enzymes of **FIG. 10**, which were added to *T. reesei* integrated strain H3A, in accordance with Example 2.

[0084] **FIGs. 12A-12B:** **FIG. 12A** depicts the expression cassette *Peg11-eg4-sucA*, as described in Example 3; **FIG. 12B** depicts the plasmid map of pCR Blunt II TOPO containing expression cassette pEG1-EG4-sucA, as described in Example 3.

[0085] **FIG. 13:** depicts the amount or percentage of glucan and xylan conversion to cellobiose, glucose, xylobiose and xylose by an enzyme composition comprising enzymes produced by the *T. reesei* integrated strain H3A transformants expressing *T. reesei* Eg4, in accordance with Example 3.

[0086] **FIG. 14:** depicts the increased percent glucan conversion observed using an increasing amount of an enzyme composition produced by H3A transformants expressing *T. reesei* Eg4. The experimental details are described in Example 3.

[0087] **FIG. 15:** provides a *T. reesei* Eg4 dosing chart for Example 4 (experiment 1). The sample “#27” is an H3A/Eg4 integrated strain as described in Example 4. The amounts of purified *T. reesei* Eg4 that were added were listed under “Sample Description” either by wt.% or by mass (in mg protein/g G+X).

[0088] **FIGs. 16A-16B:** **FIG. 16A** depicts the effect of *T. reesei* Eg4 on glucose release in saccharification of dilute ammonia pretreated corncob according to Example 4. **FIG. 16B** depicts the effect of *T. reesei* Eg4 on xylose release in saccharification of dilute ammonia pretreated corncob. The Y-axes of these figures refer to the concentrations of glucose or xylose released in the reaction mixtures. The X axes list the names/brief descriptions of the enzyme composition samples. This is according to Example 4 (experiment 1).

[0089] **FIGs. 17A-17B:** **FIG. 17A** provides another *T. reesei* Eg4 dosing chart for Example 4 (experiment 2). The samples are described similarly to those in **FIG. 15**. The amounts of purified *T. reesei* Eg4 that were added varied by smaller increments than those of Example 4, experiment 1 (above). **FIG. 17B** provides another *T. reesei* Eg4 dosing chart for Example 4 (experiment 3). The samples are described similarly to those in **FIGs. 16** and **17A**. The amounts of purified *T. reesei* Eg4 that were added varied by even finer increments than those of Example 4, experiments 1 and 2 (above)

[0090] **FIGs. 18A-18B:** **FIG. 18A** depicts the effect of *T. reesei* Eg4 in various amounts (0.05 mg/g to 1.0 mg/g) on glucose release from saccharification of dilute ammonia pretreated corncob, as described in Example 4. **FIG. 18B** depicts the effect of *T. reesei* Eg4 in various amounts (0.1 mg/g to 0.5 mg/g) on glucose release from saccharification of dilute ammonia pretreated corncob, as described in Example 4.

[0091] FIG. 19: depicts the effect of *T. reesei* Eg4 in an enzyme composition on glucose/xylose release from saccharification of different solid loadings of dilute ammonia pretreated corn stover, as described in Example 5. The solid loading is listed on the x-axis as #%.

[0092] FIG. 20: provides percentage yield of xylose monomers released from dilute ammonia pretreated corncob using an enzyme composition comprising *T. reesei* Eg4, in accordance with Example 6.

[0093] FIG. 21: provides percentage yield of glucose monomer released from dilute ammonia pretreated corncob using an enzyme composition comprising *T. reesei* Eg4, in accordance with Example 6.

[0094] FIG. 22: provides yield (mg/ml) of total fermentable monomers released from dilute ammonia pretreated corncob using an enzyme composition comprising *T. reesei* Eg4, in accordance with Example 6.

[0095] FIG. 23: compares the amounts of glucose released as a result of hydrolysis by an enzyme composition without *T. reesei* Eg4 vs. one comprising *T. reesei* Eg4 at 0.53 mg/g.

The experiment is described in Example 7.

[0100] FIG. 24: depicts the glucose monomer release as a result of treating ammonia pretreated corncob using purified *T. reesei* Eg4 alone, according to Example 7.

[0101] FIG. 25: depicts and compares the saccharification performance of the enzyme compositions produced by the *T. reesei* integrated strain H3A and the integrated strain H3A/Eg4 (strain #27), at an enzyme dosage of 14 mg/g. This is according to the description of Example 8.

[0102] FIG. 26: depicts the saccharification performance of the enzyme compositions produced by the *T. reesei* integrated strain H3A and the integrated strain H3A/Eg4 (strain #27), at various enzyme dosages, on acid pretreated corn stover. This is according to the description of Example 9.

[0103] FIG. 27: depicts the saccharification performance of the enzyme compositions produced by the *T. reesei* integrated strain H3A and the integrated strain H3A/Eg4 (strain #27) on dilute ammonia pretreated corn leaves, stalks, or cobs, according to Example 10.

[0104] FIG. 28: compares saccharification performance, in terms the amounts of glucose or xylose released, of enzyme compositions produced by the *T. reesei* integrated strain H3A and the integrated strain H3A/Eg4 (strain #27). This is according to Example 11.

[0105] FIG. 29: depicts the change in percent glucan and xylan conversion at increasing

amounts of an enzyme composition produced by the *T. reesei* integrated strain H3A/Eg4 (strain #27). This is in accordance with the description of Example 12.

[0106] FIG. 30: is a table listing the effect of *T. reesei* Eg4 addition on dilute ammonia pretreated corncob saccharification. Experimental conditions are described in Example 13.

5 [0107] FIG. 31: depicts CMC hydrolysis by *T. reesei* Eg4. Experimental conditions are described in Example 13.

[0108] FIG. 32: depicts cellobiose hydrolysis by *T. reesei* Eg4. Experimental conditions are described in Example 13.

[0109] FIG. 33: depicts amounts for various enzyme compositions for saccharification.
10 Experimental conditions are described in Example 14.

[0110] FIG. 34: depicts the amount of glucose, glucose + cellobiose, or xylose produced with each enzyme composition corresponding to FIG. 33. Experimental conditions are described in Example 14.

[0111] FIG. 35: depicts various ratios of CBH1, CBH2 and *T. reesei* Eg2 mixtures, as
15 described in Example 15.

[0112] FIG. 36: depicts glucan conversion (%) using various enzyme compositions. Experimental conditions are described in Example 15.

[0113] FIG. 37 depicts the effect of ascorbic acid when a composition comprising *T. reesei* Eg4 is used to treat Avicel in the presence or absence of CBH I, according to Example 22.

20 [0114] FIG. 38: depicts the effect of ascorbic acid on a composition comprising *T. reesei* Eg4 is used to treat Avicel in the presence/absence of CBH II, according to Example 22

[0115] FIGs. 39A-39B: FIG. 39A depicts the amount of substrate and various enzymes used in the experiment of Example 22, with the result depicted in FIG. 37. FIG. 39B depicts the amount of substrate and various enzymes used in the experiment of Example 22, with the
25 result depicted in FIG. 38.

[0116] FIG. 40: depicts glucose production from corncob hydrolysis using various enzyme compositions, in accordance with the experiments described in Example 16.

[0117] FIG. 41: depicts xylose production from corncob hydrolysis using various enzyme compositions in accordance with the description of Example 16.

30 [0118] FIG. 42: depicts viscosity of saccharification mixture using H3A and H3A added with purified Eg4 over time in accordance with the description of Example 17.

[0119] FIG. 43: depicts viscosity of saccharification mixture using H3A and H3A/Eg4#27

over time in accordance with the description of Example 18.

[0120] **FIG. 44:** depicts viscosity of saccharification of dilute ammonia pretreated corncob at 25% and 30% solids, using fermentation broths of H3A or of H3A/Eg4#27 broth at 14 mg/g cellulose, in accordance with the description of Example 19.

5 [0121] **FIG. 45:** depicts glucose concentration in 6-h saccharification, 25% dry matter, 50°C, pH5.0 using various enzyme compositions according to Example 20.

[0122] **FIG. 46:** depicts glucose concentration in 24-hour saccharification, 25% dry matter, 50°C, pH5.0 using various enzyme compositions according to Example 20.

10 [0123] **FIG. 47:** depicts glucose concentration in saccharification over time, 25% dry matter, 50°C, pH5.0 using various enzyme compositions according to Example 20.

[0124] **FIG. 48:** depicts glucan conversion in saccharification over time, 25% dry matter, 50°C, pH5.0 using various enzyme compositions according to Example 20.

[0125] **FIG. 49** provides a summary of the sequence identifies in the present disclosure.

15 [0126] **FIGs. 50A-50B:** **FIG. 50A** depicts nucleotide sequence encoding Fv3A (SEQ ID NO:35). **FIG. 50B** depicts Fv3A amino acid sequence (SEQ ID NO:36). The predicted signal sequence is underlined, and the predicted conserved domain is in bold.

[0127] **FIGs. 51A-51B:** **FIG. 51A** depicts nucleotide sequence encoding Pf43A (SEQ ID NO:37). **FIG. 51B** depicts Pf43A amino acid sequence (SEQ ID NO:38). The predicted signal sequence is underlined, the predicted conserved domain is in bold, the predicted carbohydrate binding module ("CBM") is in uppercase, and the predicted linker separating the CD and CBM is in italics.

20 [0128] **FIG. 52A-52B:** **FIG. 52A** depicts nucleotide sequence encoding Fv43E (SEQ ID NO:39). **FIG. 52B** depicts Fv43E amino acid sequence (SEQ ID NO:40). The predicted signal sequence is underlined, and the predicted conserved domain is in bold.

25 [0129] **FIGs. 53A-53B:** **FIG. 53A** depicts nucleotide sequence encoding Fv39A (SEQ ID NO:41). **FIG. 53B** depicts Fv39A amino acid sequence (SEQ ID NO:42). The predicted signal sequence is underlined, and the predicted conserved domain is in bold.

[0130] **FIGs. 54A-54B:** **FIG. 54A** depicts nucleotide sequence encoding Fv43A (SEQ ID NO:43). **FIG. 54B** depicts Fv43A amino acid sequence (SEQ ID NO:44). The predicted signal sequence is underlined, the predicted conserved domain in bold, the predicted CBM in uppercase, and the predicted linker connecting the conserved domain and CBM in italics.

30 [0131] **FIGs. 55A-55B:** **FIG. 55A** depicts nucleotide sequence encoding Fv43B (SEQ ID

NO:45). **FIG. 55B** depicts Fv43B amino acid sequence (SEQ ID NO:46). The predicted signal sequence is underlined. The predicted conserved domain is in boldface type.

[0132] FIGs. 56A-56B: FIG. 56A depicts nucleotide sequence encoding Pa51A (SEQ ID NO:47). **FIG. 56B** depicts Pa51A amino acid sequence (SEQ ID NO:48). The predicted signal sequence is underlined. The predicted L- α -arabinofuranosidase conserved domain is in bold. For expression in *T. reesei*, the genomic DNA was codon optimized (see **FIG. 73C**).

[0133] FIGs. 57A-57B: FIG. 57A depicts nucleotide sequence encoding Gz43A (SEQ ID NO:49). **FIG. 57B** depicts Gz43A amino acid sequence (SEQ ID NO:50). The predicted signal sequence is underlined, and the predicted conserved domain is in bold. For expression in *T. reesei*, the predicted signal sequence was replaced by *T. reesei* CBH1 signal sequence (myrklavisafatara (SEQ ID NO: 120)).

[0134] FIGs. 58A-58B: FIG. 58A depicts nucleotide sequence encoding Fo43A (SEQ ID NO:51). **FIG. 58B** depicts Fo43A amino acid sequence (SEQ ID NO:52). The predicted signal sequence is underlined, and the predicted conserved domain is in bold. For expression in *T. reesei*, the predicted signal sequence was replaced by *T. reesei* CBH1 signal sequence (myrklavisafatara (SEQ ID NO:120))

[0135] FIGs. 59A-59B: FIG. 59A depicts nucleotide sequence encoding Af43A (SEQ ID NO:53). **FIG. 59B** depicts Af43A amino acid sequence (SEQ ID NO:54). The predicted conserved domain is in bold.

[0136] FIGs. 60A-60B: FIG. 60A depicts nucleotide sequence encoding Pf51A (SEQ ID NO:55). **FIG. 60B** depicts Pf51A amino acid sequence (SEQ ID NO:56). The predicted signal sequence is underlined, and the predicted L- α -arabinofuranosidase conserved domain in bold. For expression in *T. reesei*, the predicted signal sequence was replaced by a codon optimized the *T. reesei* CBH1 signal sequence (myrklavisafatara (SEQ ID NO:120))

(underlined) and the Pf51A nucleotide sequence was codon optimized for expression.

[0137] FIGs. 61A-61B: FIG. 61A depicts nucleotide sequence encoding AfuXyn2 (SEQ ID NO:57). **FIG. 61B** depicts AfuXyn2 amino acid sequence (SEQ ID NO:58). The predicted signal sequence is underlined, and the predicted GH11 conserved domain in bold.

[0138] FIGs. 62A-62B: FIG. 62A depicts nucleotide sequence encoding AfuXyn5 (SEQ ID NO:59). **FIG. 62B** depicts AfuXyn5 amino acid sequence (SEQ ID NO:60). The predicted signal sequence is underlined, and the predicted GH11 conserved domain in bold.

FIGs. 63A-63B: FIG. 63A depicts nucleotide sequence encoding Fv43D (SEQ ID NO:61).

FIG. 63B depicts Fv43D amino acid sequence (SEQ ID NO:62). The predicted signal sequence is underlined. The predicted conserved domain is in bold.

[0139] FIGs. 64A-64B: FIG. 64A depicts nucleotide sequence encoding Pf43B (SEQ ID NO:63). FIG. 64B depicts Pf43B amino acid sequence (SEQ ID NO:64). The predicted
 5 signal sequence is underlined, and the predicted conserved domain is in bold.

[0140] FIGs. 65A-65B: FIG. 65A depicts nucleotide sequence encoding Fv51A (SEQ ID NO:65). FIG. 65B depicts Fv51A amino acid sequence (SEQ ID NO:66). The predicted
 signal sequence is underlined, and the predicted L- α -arabinofuranosidase conserved domain is in bold.

10 **[0141] FIGs. 66A-66B: FIG. 66A depicts nucleotide sequence encoding Cg51B (SEQ ID NO:67). FIG. 66B depicts Cg51B amino acid sequence (SEQ ID NO:68). The predicted**
 signal sequence corresponding is underlined, and the predicted conserved domain is in bold.

[0142] FIGs. 67A-67B: FIG. 67A depicts nucleotide sequence encoding Fv43C (SEQ ID NO:69). FIG. 67B depicts Fv43C amino acid sequence (SEQ ID NO:70). The predicted
 15 signal sequence is underlined, and the predicted conserved domain is in bold.

[0143] FIGs. 68A-68B: FIG. 68A depicts nucleotide sequence encoding Fv30A (SEQ ID NO:71). FIG. 68B depicts Fv30A amino acid sequence (SEQ ID NO:72). The predicted
 signal sequence is underlined.

[0144] FIGs. 69A-69B: FIG. 69A depicts nucleotide sequence encoding Fv43F (SEQ ID NO:73). FIG. 69B depicts Fv43F amino acid sequence (SEQ ID NO:74). The predicted
 20 signal sequence is underlined.

[0145] FIGs. 70A-70B: FIG. 70A depicts nucleotide sequence encoding *T. reesei* Xyn3 (SEQ ID NO:75). FIG. 70B depicts Xyn3 amino acid sequence (SEQ ID NO:76). The
 predicted signal sequence is underlined, and the predicted conserved domain is in bold.

25 **[0146] FIGs. 71A-71B: FIG. 71A depicts amino acid sequence of *T. reesei* Xyn2 (SEQ ID NO:77). The signal sequence is underlined. The predicted conserved domain is in bold. The**
 coding sequence can be found in Törrönen *et al.* Biotechnology, 1992, 10:1461-65. **FIG. 71B**
 depicts the nucleotide sequence encoding Xyn2 (SEQ ID NO:160).

[0147] FIGs. 72A-72B: FIG. 72A depicts amino acid sequence of *T. reesei* Bxl1 (SEQ ID NO:78). The signal sequence is underlined. The predicted conserved domain is in bold. The
 30 coding sequence can be found in Margolles-Clark *et al.* Appl. Environ. Microbiol. 1996, 62(10):3840-46. **FIG. 72B** depicts nucleotide sequence encoding Bxl1 (SEQ ID NO: 159)

[0148] FIGs. 73A-73F: FIG. 73A depicts amino acid sequence of *T. reesei* Bgl1 (SEQ ID NO:79). The signal sequence is underlined. The predicted conserved domain is in bold. The coding sequence can be found in Barnett *et al.* Bio-Technology, 1991, 9(6):562-567. **FIG.**

73B depicts deduced cDNA for Pa51A (SEQ ID NO:80). **FIG. 73C** depicts codon optimized cDNA for Pa51A (SEQ ID NO:81). **FIG. 73D:** depicts coding sequence for a construct comprising a CBH1 signal sequence (underlined) upstream of genomic DNA encoding mature Gz43A (SEQ ID NO:82). **FIG. 73E:** depicts coding sequence for a construct comprising a CBH1 signal sequence (underlined) upstream of genomic DNA encoding mature Fo43A (SEQ ID NO:83). **FIG. 73F:** depicts codon optimized coding sequence for a construct comprising a CBH1 signal sequence (underlined) upstream of codon optimized DNA encoding mature Pf51A (SEQ ID NO:92).

[0149] FIGs. 74A-74B: FIG. 74A depicts nucleotide sequence encoding Pa3D (SEQ ID NO:93). **FIG. 74B** depicts amino acid sequence of Pa3D (SEQ ID NO:94). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0150] FIGs. 75A-75B: FIG. 75A depicts nucleotide sequence encoding Fv3G (SEQ ID NO:95). **FIG. 75B** depicts amino acid sequence of Fv3G (SEQ ID NO:96). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0151] FIGs. 76A-76B: FIG. 76A depicts nucleotide sequence encoding Fv3D (SEQ ID NO:97). **FIG. 76B** depicts amino acid sequence of Fv3D (SEQ ID NO:98). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0152] FIGs. 77A-77B: FIG. 77A depicts nucleotide sequence encoding Fv3C (SEQ ID NO:99). **FIG. 77B** depicts amino acid sequence of Fv3C (SEQ ID NO:100). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0153] FIGs. 78A-78B: FIG. 78A depicts nucleotide sequence encoding Tr3A (SEQ ID NO:101). **FIG. 78B** depicts amino acid sequence of Tr3A (SEQ ID NO:102). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0154] FIGs. 79A-79B: FIG. 79A depicts nucleotide sequence encoding Tr3B (SEQ ID NO:103). **FIG. 79B** depicts amino acid sequence of Tr3B (SEQ ID NO:104). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0155] FIGs. 80A-80B: FIG. 80A depicts nucleotide sequence encoding Te3A (SEQ ID NO:105). **FIG. 80B** depicts amino acid sequence of Te3A (SEQ ID NO:106). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0156] **FIGs. 81A-81B: FIG. 81A** depicts nucleotide sequence encoding An3A (SEQ ID NO:107). **FIG. 81B** depicts amino acid sequence of An3A (SEQ ID NO:108). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0157] **FIGs. 82A-82B: FIG. 82A** depicts nucleotide sequence encoding Fo3A (SEQ ID NO:109). **FIG. 82B** depicts amino acid sequence of Fo3A (SEQ ID NO:110). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0158] **FIGs. 83A-83B: FIG. 83A** depicts nucleotide sequence encoding Gz3A (SEQ ID NO:111). **FIG. 83B** depicts amino acid sequence of Gz3A (SEQ ID NO:112). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0159] **FIGs. 84A-84B: FIG. 84A** depicts nucleotide sequence encoding Nh3A (SEQ ID NO:113). **FIG. 84B** depicts amino acid sequence of Nh3A (SEQ ID NO:114). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0160] **FIGs. 85A-85B: FIG. 85A** depicts nucleotide sequence encoding Vd3A (SEQ ID NO:115). **FIG. 85B** depicts amino acid sequence of Vd3A (SEQ ID NO:116). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0161] **FIGs. 86A-86B: FIG. 86A** depicts nucleotide sequence encoding Pa3G (SEQ ID NO:117). **FIG. 86B** depicts amino acid sequence of Pa3G (SEQ ID NO:118). The predicted signal sequence is underlined, and the predicted conserved domains are in bold.

[0162] **FIG. 87:** depicts amino acid sequence encoding Tn3B (SEQ ID NO:119). The standard signal prediction program, Signal P (www.cbs.dtu.dk/services/SignalP/) provided no predicted signal.

[0163] **FIG. 88:** depicts a partial amino acid sequence alignment of the CBM domains of *T. reesei* Eg4 (SEQ ID NO:27) with Tr6A (SEQ ID NO:31) and with Tr7A (SEQ ID NO:32).

[0164] **FIGs. 89A-89C: FIG. 89A** depicts amino acid sequence of Eg6 (SEQ ID NO:33)

from *T. reesei*. The bolded amino acid sequence is the predicted signal peptide sequence.

FIG. 89B depicts amino acid sequence of *S. coccosporum* endoglucanase SEQ ID NO:34;

FIG. 89C depicts the nucleotide sequence encoding a GH61A from *Thermoascus aurantiacus*, SEQ ID NO:149.

[0165] **FIGs 90A-90I: FIG. 90A** depicts amino acid sequence of Afu7A (SEQ ID NO:150), a homolog of CBH1 of *T. reesei*. **FIG. 90B** depicts amino acid sequence of Afu7B (SEQ ID NO:151), a homolog of CBH1 of *T. reesei*. **FIG. 90C** depicts amino acid sequence of Cg7A (SEQ ID NO:152), a homolog of CBH1 of *T. reesei*. **FIG. 90D** depicts amino acid

sequence of Cg7B(SEQ ID NO:153), a homolog of CBH1 of *T. reesei*. **FIG. 90E** depicts amino acid sequence of Tt7A(SEQ ID NO:154), a homolog of CBH1 of *T. reesei*. **FIG. 90F** depicts amino acid sequence of Tt7B(SEQ ID NO:155), a homolog of CBH1 of *T. reesei*. **FIG. 90G** depicts amino acid sequence of St6A (SEQ ID NO:156), a homolog of CBH2 of *T. reesei*. **FIG. 90H** depicts amino acid sequence of St6B (SEQ ID NO:157), a homolog of CBH2 of *T. reesei*. **FIG. 90I** amino acid sequence of Tt6A (SEQ ID NO:158), a homolog of CBH2 of *T. reesei*.

DETAILED DESCRIPTION OF THE INVENTION

[0166] Unless defined otherwise, all technical and scientific terms used herein have the meaning as commonly understood by a skilled person in the art to which this invention belongs. Singleton, et al., **DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY**, 2D ED., John Wiley and Sons, New York (1994), and Hale & Marham, **THE HARPER COLLINS DICTIONARY OF BIOLOGY**, Harper Perennial, N.Y. (1991) provide one of skill with a general dictionary of many of the terms used in this invention. Although any methods and materials similar or equivalent to those described herein can be used in the practice of the present invention, the preferred methods and materials are described. Numeric ranges are inclusive of the numbers defining the range. The invention is not limited to the particular methodology, protocols, and reagents described, as these may vary.

[0167] The headings provided herein do not limit the various aspects or embodiments of the invention that can be had by reference to the specification as a whole. Accordingly the terms defined below are more fully defined by reference to the specification as a whole.

[0168] The present disclosure provides compositions comprising a polypeptide having glycosyl hydrolase family 61 (“GH61”)/endoglucanase activity, polypeptides having GH61/endoglucanase activity, nucleotides encoding a polypeptide provided herein, vectors containing nucleotide provided herein, and cells containing nucleotide and/or vector provided herein. The present disclosure further provides methods of hydrolyzing a biomass material and methods of reducing the viscosity of a biomass-containing mixture using a composition provided herein.

[0169] The term “isolated” as used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules separated from other DNAs or RNAs, respectively, which are present in the natural source of the nucleic acid. Moreover, by an “isolated nucleic acid” is

meant to include nucleic acid fragments, which are not naturally occurring as fragments and would not be found in the natural state. The term “isolated” is also used herein to refer to polypeptides, which are isolated from other cellular proteins and is meant to encompass both purified and recombinant polypeptides. The term “isolated” as used herein also refers to a nucleic acid or polypeptide that may be substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques. The term “isolated” as used herein additionally refers to a nucleic acid or polypeptide that may be substantially free of chemical precursors or other chemicals when chemically synthesized.

[0170] As used herein, a “variant” of polypeptide X refers to a polypeptide having the amino acid sequence of polypeptide X with one or more altered amino acid residues. The variant may have conservative or nonconservative changes. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without affecting biological activity may be found using computer programs known in the art, e.g., LASERGENE software (DNASTAR). A variant of the invention includes polypeptides comprising altered amino acid sequences in comparison with a precursor enzyme amino acid sequence, wherein the variant enzyme retains the characteristic cellulolytic nature of the precursor enzyme but may have altered properties in some specific aspects, e.g., an increased or decreased pH optimum, an increased or decreased oxidative stability; an increased or decreased thermostability, and increased or decreased level of specific activity towards one or more substrates, as compared to the precursor enzyme.

[0171] As used herein, a polypeptide or nucleic acid that is “heterologous” to a host cell refers to a polypeptide or nucleic acid that does not naturally occur in a host cell.

[0172] Reference to “about” a value or parameter herein includes (and describes) variations that are directed to that value or parameter per se. For example, description referring to “about X” includes description of “X”.

[0173] As used herein and in the appended claims, the singular forms “a,” “or,” and “the” include plural referents unless the context clearly dictates otherwise.

[0174] It is understood that aspects and variations of the methods and compositions described herein include “consisting” and/or “consisting essentially of” aspects and variations.

Polypeptides

[0175] The disclosure provides polypeptides (*e.g.*, isolated, synthetic, or recombinant

polypeptides) having GH61/endoglucanase activity. For example, the present disclosure provides GH61 endoglucanases from various species or variants thereof, endoglucanase IV (or endoglucanase 4) polypeptides (also described herein as “Eg4” or “EG4”, which are used interchangeably herein) from various species or variants thereof, and *Trichoderma reesei* Eg4 polypeptide or variants thereof. In some aspects, the polypeptide is isolated.

Glycoside hydrolase family 61 (“GH61”) enzymes

[0176] Glycoside hydrolase family 61 (“GH61”) enzymes have been identified in Eukaryota. A weak endoglucanase activity has been observed for Cel61A from *Hypocrea jecorina* (Karlsson et al, Eur J Biochem, 2001, 268(24):6498-6507), which is thus said to have GH61/endoglucanase activity. GH61 polypeptides potentiate enzymatic hydrolysis of lignocellulosic substrates by cellulases (Harris et al, 2010, Biochemistry, 49(15) 3305-16). Studies on homologous polypeptides involved in chitin degradation predict that GH61 polypeptides may employ an oxidative hydrolysis mechanism that requires an electron donor substrate and in which divalent metal ions are involved (Vaaje-Kolstad, 2010, Science, 330(6001), 219-22). This agrees with the observation that the synergistic effect of GH61 polypeptides on lignocellulosic substrate degradation is dependent on divalent ions (Harris et al, 2010, Biochemistry, 49(15) 3305-16). A number of available structures of GH61 polypeptides have divalent atoms bound by a number of conserved amino acid residues (Karkehabadi, 2008, J. Mol. Biol., 383(1) 144-54; Harris et al, 2010, Biochemistry, 49(15) 3305-16). It has been reported that the GH61 polypeptides have a flat surface at the metal binding site that is formed by conserved residues and might be involved in substrate binding (Karkehabadi, 2008, J. Mol. Biol., 383(1), 144-54).

[0177] The present disclosure provides polypeptides having GH61/endoglucanase activity (e.g., isolated polypeptide) which can be a GH61 endoglucanase or endoglucanase IV (“EG IV”) from various species, or can also be a polypeptide from various species corresponding to (sharing homology with, sharing functional domains, sharing GH61 motif(s), and/or sharing conservative residues with) a GH61 endoglucanase (e.g., a *Trichoderma reesei* Eg4 polypeptide). Such species include *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, *Chrysosporium*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *Fusarium bactridioides*, *Fusarium cerealis*,

- Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium gramineum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*,
- 5 *Fusarium venenatum*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Neurospora intermedia*, *Penicillium purpurogenum*,
- 10 *Penicillium canescens*, *Penicillium solitum*, *Penicillium funiculosum* *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, *Trichoderma viride*, *Geosmithia emersonii*, or *G. stearothermophilus*.
- 15 **[0178]** Polypeptides having GH61/endoglucanase activity include a number of GH61 endoglucanases listed in FIG. 1. For example, suitable GH61 endoglucanases include those comprising amino acid sequences that are at least about 60% identical to the various sequences listed in FIG. 1, including, for example, those represented by their GenBank Accession Numbers CAB97283.2, CAD70347.1, CAD21296.1, CAE81966.1, CAF05857.1,
- 20 EAA26873.1, EAA29132.1, EAA30263.1, EAA33178.1, EAA33408.1, EAA34466.1, EAA36362.1, EAA29018.1, and EAA29347.1, or St61 from *S. thermophilum* 24630, St61A from *S. thermophilum* 23839c, St61B from *S. thermophilum* 46583, St61D from *S. thermophilum* 80312, Afu61a from *A. fumigatus* Afu3g03870 (NCBI Ref: XP_748707), an endoglucanase having NCBI Ref: XP_750843.1 from *A. fumigatus* Afu6g09540, an
- 25 endoglucanase from *A. fumigatus* EDP47167, an endoglucanase from *T. terrestris* 16380, an endoglucanase from *T. terrestris* 155418, an endoglucanase from *T. terrestris* 68900, Cg61A (Accession Number EAQ86340.1) from *C. globosum*, *T. reesei* Eg7, *T. reesei* Eg4, and an endoglucanase with GenBank Accession Number XP_752040 from *A. fumigatus* Af293. In some aspects, a suitable GH61 endoglucanase polypeptide of the invention comprises an
- 30 amino acid sequence of at least about 60% (e.g., at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) sequence identity to any one of SEQ ID NOs: 1-29 and 148. In some aspects, a suitable GH61 endoglucanase polypeptide

of the invention comprises one or more of the amino acid sequence motifs selected from: (1) SEQ ID NOs:84 and 88; (2) SEQ ID NOs:85 and 88; (3) SEQ ID NO:86; (4) SEQ ID NO:87; (5) SEQ ID NOs:84, 88 and 89; (6) SEQ ID NOs:85, 88, and 89; (7) SEQ ID NOs: 84, 88, and 90; (8) SEQ ID NOs: 85, 88 and 90; (9) SEQ ID NOs:84, 88 and 91; (10) SEQ ID NOs: 85, 88 and 91; (11) SEQ ID NOs: 84, 88, 89 and 91; (12) SEQ ID NOs: 84, 88, 90 and 91; (13) SEQ ID NOs: 85, 88, 89 and 91; and (14) SEQ ID NOs: 85, 88, 90 and 91. The polypeptide may be at least 100 (e.g., 110, 120, 130, 140, 150, 160, 170, 180, 200, 220, 250 or more) residues in length.

[0179] Polypeptides having GH61/endoglucanase activity (e.g., isolated polypeptide)

provided herein may also be a variant of a GH61 endoglucanase, e.g., any of the polypeptides with amino acid sequences shown FIG. 1 of the present disclosure. For example, suitable GH61 endoglucanases include those represented by their GenBank Accession Numbers CAB97283.2, CAD70347.1, CAD21296.1, CAE81966.1, CAF05857.1, EAA26873.1, EAA29132.1, EAA30263.1, EAA33178.1, EAA33408.1, EAA34466.1, EAA36362.1, EAA29018.1, and EAA29347.1, or St61 from *S. thermophilum* 24630, St61A from *S. thermophilum* 23839c, St61B from *S. thermophilum* 46583, St61D from *S. thermophilum* 80312, Afu61a from *A. fumigatus* Afu3g03870 (NCBI Ref: XP_748707), an endoglucanase with NCBI Ref: XP_750843.1 from *A. fumigatus* Afu6g09540, an endoglucanase from *A. fumigatus* EDP47167, an endoglucanase from *T. terrestris* 16380, an endoglucanase from *T. terrestris* 155418, an endoglucanase from *T. terrestris* 68900, Cg61A (EAQ86340.1) from *C. globosum*, *T. reesei* Eg7, *T. reesei* Eg4, and an endoglucanase with GenBank Accession: XP_752040 from *A. fumigatus* Af293. In some aspects, the polypeptide having GH61/endoglucanase activity (e.g., isolated polypeptide) is a variant of EG IV. In some aspects, the polypeptide having GH61/ endoglucanase activity (e.g., isolated polypeptide) is a variant of a GH61 endoglucanase, wherein the variant has an amino acid sequence having at least about 60% (e.g., at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%) identity as any one of the amino acid sequences SEQ ID NOs: 1-29 and 148.

[0180] An alignment using amino acid sequences SEQ ID NOs:1-29 and 148 was performed and the alignment result is shown in FIG. 3. FIG.2 shows the percent identity and divergence results from comparison of the amino acid sequences of the polypeptides. The alignment indicated that the GH61 endoglucanase polypeptides share certain sequence motifs, and such motifs are shown in FIG. 7 of the present disclosure.

[0181] Accordingly, the present disclosure provides polypeptides (*e.g.*, isolated, synthetic, or recombinant polypeptides) having GH61/endoglucanase activity, which may be a GH61 endoglucanase or a variant thereof, and the variant may comprise at least one motif (at least any of 2, 3, 4, 5, 6, 7, or 8) selected from SEQ ID NOs:84-91. Each of the “a”s in sequence motifs with SEQ ID NOs:84-91 (described in FIG.7) represents an amino acid that may be any one of alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, or valine. For example, in some aspects, the disclosure provides polypeptides (*e.g.*, isolated, synthetic, or recombinant polypeptides) comprising at least one sequence motif, such as at least one (*e.g.*, 2, 3, 4, 5, 6, 7, or 8) of SEQ ID NOs: 84, 85, 86, 87, 88, 89, 90, and 91. In some aspects, the disclosure provides polypeptides (*e.g.*, isolated, synthetic, or recombinant polypeptides) comprising one or more of the sequence motifs selected from the group consisting of: (1) SEQ ID NOs:84 and 88; (2) SEQ ID NOs:85 and 88; (3) SEQ ID NO:86; (4) SEQ ID NO:87; (5) SEQ ID NOs:84, 88 and 89; (6) SEQ ID NOs:85, 88, and 89; (7) SEQ ID NOs: 84, 88, and 90; (8) SEQ ID NOs: 85, 88 and 90; (9) SEQ ID NOs:84, 88 and 91; (10) SEQ ID NOs: 85, 88 and 91; (11) SEQ ID NOs: 84, 88, 89 and 91; (12) SEQ ID NOs: 84, 88, 90 and 91; (13) SEQ ID NOs: 85, 88, 89 and 91; and (14) SEQ ID NOs: 85, 88, 90 and 91, , over a region of at least about 10, *e.g.*, at least about any of 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, or 350 residues, or over the full length of the immature polypeptide, the full length mature polypeptide, the full length of the conserved domain, and/or the full length CBM. The conserved domain can be a predicted catalytic domain (“CD”). Exemplary polypeptides also include fragments of at least about 10, *e.g.*, at least about any of 15, 20, 25, 30, 35, 40, 45, 50, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, or 600 residues in length. The fragments can comprise a conserved domain and/or a CBM. Where a fragment comprises a conserved domain and a CBM of an enzyme, the fragment optionally includes a linker separating the two. The linker can be a native linker or a heterologous linker. In some aspects, the polypeptide has GH61/endoglucanase activity.

[0182] In some aspects, the polypeptide having GH61/endoglucanase activity is a GH61 endoglucanase or a variant thereof, an enzyme comprising any one of SEQ ID NOs: 1-29 and 148, or a variant thereof, an EG IV or a variant thereof, or a *T. reesei* Eg4 or a variant thereof. A variant described here has endoglucanase activity. The polypeptide having GH61/

endoglucanase activity (including a variant) may comprise a CBM domain (e.g., functional CBM domain). The polypeptide having GH61/ endoglucanase activity (including a variant) may comprise a catalytic domain (e.g., function catalytic domain).

[0183] *T. reesei* Eg4 is a GH61 endoglucanase polypeptide. The amino acid sequence of *T. reesei* Eg4 (SEQ ID NO:27) is shown in **FIGs. 1, 4B and 5**. SEQ ID NO:27 is the sequence of the immature *T. reesei* Eg4. *T. reesei* Eg4 has a predicted signal sequence corresponding to residues 1 to 21 of SEQ ID NO:27 (underlined); cleavage of the signal sequence is predicted to yield a mature polypeptide having a sequence corresponding to residues 22 to 344 of SEQ ID NO:27. The predicted conserved domains correspond to residues 22-256 and 307-343 of SEQ ID NO:27, with the latter being the predicted carbohydrate-binding domain (CBM). *T. reesei* Eg4 was shown to have endoglucanase activity in, for example, an enzymatic assay using carboxy methyl cellulose as substrates. Methods of measuring endoglucanase activity are also known to one skilled in the art.

[0184] The disclosure further provides a variant of *Trichoderma reesei* Eg4 polypeptide, which may comprise a sequence having at least about 60% (e.g., at least about 65%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%) sequence identity to at least about 50 (e.g., at least about 55, 60, 65, 70, 75, 100, 125, 150, 175, 200, 250, or 300) contiguous amino acid residues among residues 22 to 344 of SEQ ID NO:27. For example, the disclosure provides variants of *T. reesei* Eg4 polypeptide. Such variants may have at least about 70% (e.g., at least about 70%, 75%, 80%, 85%, 88%, 90%, 92.5%, 95%, 96%, 97%, 98%, or 99%) identity to residues 22 to 344 of SEQ ID NO:27. The polypeptide or a variant thereof may be isolated. The polypeptide or a variant thereof may have endoglucanase activity.

[0185] *T. reesei* Eg4 residues H22, H107, H184, Q193, and Y195 were predicted to function as metal coordinator residues; residues D61 and G63 were predicted to be conserved surface residues; and residue Y232 were predicted to be involved in activity, based on an amino acid sequence alignment of a number of known endoglucanases, e.g., an endoglucanase from *T. terrestris* (Accession No. ACE10234, also termed “TtEG” herein) (SEQ ID NO:29), and another endoglucanase Eg7 (Accession No. ADA26043.1) from *T. reesei* (also termed “TrEGb” or “TrEG7” herein), with *T. reesei* Eg4 (see, **FIG. 5**). The predicted conserved residues in *T. reesei* Eg4 A are shown in **FIGs. 6A and 6B**. A variant of

T. reesei Eg4 polypeptide may be unaltered, as compared to a native *T. reesei* Eg4, at residues H22, H107, H184, Q193, Y195, D61, G63, and Y232. A variant of *T. reesei* Eg4 polypeptide may be unaltered in at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among TrEGb, TtEG, and *T. reesei* Eg4, as shown in the alignment of **FIG. 5**. A variant of *T. reesei* Eg4 polypeptide may comprise the entire predicted conserved domains of native *T. reesei* Eg4. See **FIGs. 5** and **6**. An exemplary variant of *T. reesei* Eg4 polypeptide comprises a sequence having at least about any of 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature *T. reesei* Eg4 sequence shown in **FIG. 4B** (e.g., residues 22 to 344 of SEQ ID NO:27). In some aspects, the variant of *T. reesei* Eg4 polypeptide has endoglucanase (e.g., endoglucanase IV (EGIV)) activity.

[0186] In some aspects, a variant of *T. reesei* Eg4 polypeptide has endoglucanase activity and comprises an amino acid sequence with at least about any of 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:27, or to residues (i) 22-255, (ii) 22-343, (iii) 307-343, (iv) 307-344, or (v) 22-344 of SEQ ID NO:27.

[0187] In some aspects, the polypeptide or a variant thereof comprises residues corresponding to at least about 3 residues (e.g., at least about any of 4, 5, 6, 7, 8, 9, 10, 11, or 12) of H22, D61, G63, C77, H107, R177, E179, H184, Q193, C198, Y195, and Y232 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to H22, D61, G63, C77, H107, R177, E179, H184, Q193, C198, Y195, and Y232 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to at least 3 residues (e.g., at least about any of 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, or 19) of G313, Q314, C315, G316, G317, S321, G322, P323, T324, C326, A327, T331, C332, N336, Y338, Y339, Q341, C342, and L343 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to G313, Q314, C315, G316, G317, S321, G322, P323, T324, C326, A327, T331, C332, N336, Y338, Y339, Q341, C342, and L343 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises a CBM domain (e.g., functional CBM domain). In some aspects, the polypeptide or a variant thereof comprises a catalytic domain (e.g., functional catalytic domain). The polypeptide suitably has endoglucanase activity.

[0188] A variant of GH61 endoglucanase, an endoglucanase comprising any one of SEQ

ID NOs:1-29 and 148, an EG IV, or *Trichoderma reesei* Eg4 polypeptide may be made using amino acid substitution. Conservative substitutions are shown in the table below under the heading of “conservative substitutions”. Substitutions may also be exemplary substitution shown in the table below.

5 Table 1: Amino Acid Substitutions.

Original Residue	Conservative Substitutions	Exemplary Substitutions
Ala (A)	Val	Val; Leu; Ile
Arg (R)	Lys	Lys; Gln; Asn
Asn (N)	Gln	Gln; His; Asp, Lys; Arg
Asp (D)	Glu	Glu; Asn
Cys (C)	Ser	Ser; Ala
Gln (Q)	Asn	Asn; Glu
Glu (E)	Asp	Asp; Gln
Gly (G)	Ala	Ala
His (H)	Arg	Asn; Gln; Lys; Arg
Ile (I)	Leu	Leu; Val; Met; Ala; Phe; Norleucine
Leu (L)	Ile	Norleucine; Ile; Val; Met; Ala; Phe
Lys (K)	Arg	Arg; Gln; Asn
Met (M)	Leu	Leu; Phe; Ile
Phe (F)	Tyr	Leu; Val; Ile; Ala; Tyr
Pro (P)	Ala	Ala
Ser (S)	Thr	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr	Tyr; Phe
Tyr (Y)	Phe	Trp; Phe; Thr; Ser
Val (V)	Leu	Ile; Leu; Met; Phe; Ala; Norleucine

[0189] Substantial modifications in the enzymatic properties of the polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

- (1) Non-polar: Norleucine, Met, Ala, Val, Leu, Ile;
- (2) Polar without charge: Cys, Ser, Thr, Asn, Gln;
- (3) Acidic (negatively charged): Asp, Glu;
- (4) Basic (positively charged): Lys, Arg;
- (5) Residues that influence chain orientation: Gly, Pro; and
- (6) Aromatic: Trp, Tyr, Phe, His.

[0190] Non-conservative substitutions are made by exchanging a member of one of these classes for another class. Any cysteine residue not involved in maintaining the proper conformation of the polypeptide also may be substituted, generally with serine, to improve the oxidative stability of the molecule and prevent aberrant cross-linking. Conversely, 5 cysteine bond(s) may be added to the polypeptide to improve its stability.

[0191] In some aspects, a polypeptide (*e.g.*, isolated, synthetic, or recombinant polypeptide) having GH61/endoglucanase activity is a fusion or chimeric polypeptide that includes a domain of a polypeptide of the present disclosure attached to one or more fusion segments, which are typically heterologous to the polypeptide (*e.g.*, derived from a different 10 source than the polypeptide of the disclosure). Suitable fusion or chimeric segments include, without limitation, segments that can enhance a polypeptide's stability, provide other desirable biological activity or enhanced levels of desirable biological activity, and/or facilitate purification of the polypeptide (*e.g.*, by affinity chromatography). A suitable fusion segment can be a domain of any size that has the desired function (*e.g.*, imparts increased 15 stability, solubility, action or biological activity; and/or simplifies purification of a polypeptide). A fusion or hybrid polypeptide of the invention can be constructed from two or more fusion or chimeric segments, each of which or at least two of which are derived from a different source or microorganism. Fusion or hybrid segments can be joined to amino and/or carboxyl termini of the domain(s) of a polypeptide of the present disclosure. The fusion 20 segments can be susceptible to cleavage. There may be some advantage in having this susceptibility, for example, it may enable straight-forward recovery of the polypeptide of interest. Fusion polypeptides may be produced by culturing a recombinant cell transfected with a fusion nucleic acid that encodes a polypeptide, which includes a fusion segment attached to either the carboxyl or amino terminal end, or fusion segments attached to both the 25 carboxyl and amino terminal ends, of a polypeptide, or a domain thereof.

[0192] Accordingly, polypeptides of the present disclosure also include expression products of gene fusions (*e.g.*, an overexpressed, soluble, and active form of expression product), of mutagenized genes (*e.g.*, genes having codon modifications to enhance gene transcription and translation), and of truncated genes (*e.g.*, genes having signal sequences 30 removed or substituted with a heterologous signal sequence).

[0193] Glycosyl hydrolases that utilize insoluble substrates are often modular enzymes. They may comprise catalytic modules appended to one or more non-catalytic carbohydrate-

binding domains (CBMs). In nature, CBMs are thought to promote the glycosyl hydrolase's interaction with its target substrate polysaccharide. Thus, the disclosure provides chimeric enzymes having altered substrate specificity; including, for example, chimeric enzymes having multiple substrates as a result of "spliced-in" heterologous CBMs. The heterologous

5 CBMs of the chimeric enzymes of the disclosure can also be designed to be modular, such that they are appended to a catalytic module or catalytic domain (a "CD", *e.g.*, at an active site), which can likewise be heterologous or homologous to the glycosyl hydrolase.

[0194] Thus, the disclosure provides peptides and polypeptides consisting of, or comprising, CBM/CD modules, which can be homologously paired or joined to form

10 chimeric (heterologous) CBM/CD pairs. Thus, these chimeric polypeptides/peptides can be used to improve or alter the performance of an enzyme of interest.

[0195] In some aspects, there is provided a polypeptide having GH61/endoglucanase activity, which comprises at least one CD and/or CBM of any one of the polypeptides with sequences shown in FIG 1 of the present disclosure. For example, suitable GH61

15 endoglucanase polypeptides of FIG. 1 includes those that are represented by their GenBank Accession Numbers CAB97283.2, CAD70347.1, CAD21296.1, CAE81966.1, CAF05857.1, EAA26873.1, EAA29132.1, EAA30263.1, EAA33178.1, EAA33408.1, EAA34466.1, EAA36362.1, EAA29018.1, and EAA29347.1, or St61 from *S. thermophilum* 24630, St61A from *S. thermophilum* 23839c, St61B from *S. thermophilum* 46583, St61D from *S.*

20 *thermophilum* 80312, Afu61a from *A. fumigatus* Afu3g03870 (NCBI Ref: XP_748707), an endoglucanase of NCBI Ref: XP_750843.1 from *A. fumigatus* Afu6g09540, an endoglucanase of *A. fumigatus* EDP47167, an endoglucanase of *T. terrestris* 16380, an endoglucanase of *T. terrestris* 155418, an endoglucanase of *T. terrestris* 68900, Cg61A (EAQ86340.1) from *C. globosum*, *T. reesei* Eg7, *T. reesei* Eg4, and an endoglucanase with

25 GenBank Accession: XP_752040 from *A. fumigatus* Af293. The polypeptide may suitably be a fusion polypeptide comprising functional domains from two or more different polypeptides (*e.g.*, a CBM from one polypeptide linked to a CD from another polypeptide).

[0196] The polypeptides of the disclosure can suitably be obtained and/or used in

"substantially pure" form. For example, a polypeptide of the disclosure constitutes at least

30 about 80 wt.% (*e.g.*, at least about any of 85 wt.%, 90 wt.%, 91 wt.%, 92 wt.%, 93 wt.%, 94 wt.%, 95 wt.%, 96 wt.%, 97 wt.%, 98 wt.%, or 99 wt.%) of the total protein in a given composition, which also includes other ingredients such as a buffer or solution.

[0197] Also the polypeptides of the disclosure may suitably be obtained and/or used in culture broths (*e.g.*, a filamentous fungal culture broth). The culture broth may be an engineered enzyme composition, *e.g.*, the culture broth may be produced by a recombinant host cell engineered to express a heterologous polypeptide of the disclosure, or by a recombinant host cell engineered to express an endogenous polypeptide of the disclosure in greater or lesser amounts than the endogenous expression levels (*e.g.*, in an amount that is 1-, 2-, 3-, 4-, 5-, or more- fold greater or less than the endogenous expression levels). Furthermore, the culture broths may be produced by certain “integrated” host cell strains that are engineered to express a plurality of the polypeptides of the disclosure in desired ratios.

10 *Nucleic acids, expression cassettes, vectors, and host cells*

[0198] The disclosure provides nucleic acids (*e.g.*, isolated, synthetic or recombinant nucleic acids) encoding polypeptides provided above, *e.g.*, polypeptides having GH61/endoglucanase activity, GH61 endoglucanase or a variant thereof, EG IV or a variant thereof, *T. reesei* Eg4 or a variant thereof. In certain aspects, the disclosure provides nucleic acids (*e.g.*, isolated, synthetic or recombinant nucleic acids) encoding a polypeptide comprising any one of SEQ ID NOs:1-29 and 148, or a polypeptide having at least about 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identity to any one of SEQ ID NOs: 1-29 and 148.

[0199] In certain aspects, the disclosure provides nucleic acids (*e.g.*, isolated, synthetic or recombinant nucleic acids) encoding any one of the polypeptides having GH61/endoglucanase activity (including a variant of a GH61 endoglucanase) comprising one or more sequence motif selected from: (1) SEQ ID NOs:84 and 88; (2) SEQ ID NOs:85 and 88; (3) SEQ ID NO:86; (4) SEQ ID NO:87; (5) SEQ ID NOs:84, 88 and 89; (6) SEQ ID NOs:85, 88, and 89; (7) SEQ ID NOs: 84, 88, and 90; (8) SEQ ID NOs: 85, 88 and 90; (9) SEQ ID NOs:84, 88 and 91; (10) SEQ ID NOs: 85, 88 and 91; (11) SEQ ID NOs: 84, 88, 89 and 91; (12) SEQ ID NOs: 84, 88, 90 and 91; (13) SEQ ID NOs: 85, 88, 89 and 91: and (14) SEQ ID NOs: 85, 88, 90 and 91. The disclosure further provides nucleic acids (*e.g.*, isolated, synthetic or recombinant nucleic acids) encoding a polypeptide having GH61/endoglucanase activity (including a variant of a GH61 endoglucanase) that comprises a CBM domain (*e.g.*, functional CBM domain) and/or catalytic domain (*e.g.*, functional catalytic domain).

[0200] The disclosure further provides nucleic acids (*e.g.*, isolated, synthetic or recombinant nucleic acids) encoding variants of *T. reesei* Eg4 polypeptide. Such variants may

have at least about 60% (*e.g.*, at least about any of 60%, 65%, 70%, 75%, 80%, 85%, 88%, 90%, 92.5%, 95%, 96%, 97%, 98%, or 99%) sequence identity to residues 22 to 344 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof has endoglucanase activity.

The polypeptide or a variant thereof may comprise residues corresponding to at least about 5

residues (*e.g.*, at least about any of 6, 7, 8, 9, 10, 11, or 12) of H22, D61, G63, C77, H107, R177, E179, H184, Q193, C198, Y195, and Y232 of SEQ ID NO:27. The polypeptide or a variant thereof may comprise residues corresponding to H22, D61, G63, C77, H107, R177, E179, H184, Q193, C198, Y195, and Y232 of SEQ ID NO:27. The polypeptide or a variant thereof may comprise residues corresponding to at least 5 residues (*e.g.*, at least about any of

6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, or 19) of G313, Q314, C315, G316, G317, S321, G322, P323, T324, C326, A327, T331, C332, N336, Y338, Y339, Q341, C342, and L343 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to G313, Q314, C315, G316, G317, S321, G322, P323, T324, C326, A327, T331, C332, N336, Y338, Y339, Q341, C342, and L343 of SEQ ID NO:27.

[0201] The disclosure provides nucleic acids (*e.g.*, isolated, synthetic or recombinant nucleic acids) comprising a nucleic acid sequence having at least about 70%, *e.g.*, at least about any of 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%, or complete (100%) identity to nucleic acid sequence SEQ ID NO:30, over a region of at least about 10, *e.g.*, at least about any of 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, or 1050 nucleotides. In some aspects, the disclosure provides nucleic acids encoding any one of the polypeptides provided herein. Also provided herein are isolated nucleic acids having at least about 80% (*e.g.*, at least about any of 85%, 88%, 90%, 92.5%, 95%, 96%, 97%, 98%, or 99%) identity to SEQ ID NO:30.

[0202] In some aspects, there is provided a nucleic acid (*e.g.*, isolated, synthetic or recombinant nucleic acid) encoding a polypeptide comprising an amino acid sequence with at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:27, or to residues (i) 22-255, (ii) 22-343, (iii) 307-343, (iv) 307-344, or (v) 22-344 of SEQ ID NO:27. In some aspects, there is provided a nucleic acid (*e.g.*, isolated, synthetic or recombinant nucleic acid) having at least 70% (*e.g.*, at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%,

99% or more) sequence identity to SEQ ID NO:30, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:30, or to a fragment thereof. As used herein, the term “hybridizes under low stringency, medium stringency, high stringency, or very high stringency conditions” describes conditions for hybridization and washing. Guidance for performing hybridization reactions can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1 - 6.3.6. Aqueous and nonaqueous methods are described in that reference and either method can be used. Specific hybridization conditions referred to herein are as follows: 1) low stringency hybridization conditions in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by two washes in 0.2X SSC, 0.1% SDS at least at 50°C (the temperature of the washes can be increased to 55°C for low stringency conditions); 2) medium stringency hybridization conditions in 6X SSC at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 60°C; 3) high stringency hybridization conditions in 6X SSC at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 65°C; and preferably 4) very high stringency hybridization conditions are 0.5M sodium phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2X SSC, 1% SDS at 65°C. Very high stringency conditions (4) are the preferred conditions unless otherwise specified.

[0203] The disclosure also provides expression cassettes and/or vectors comprising any of the above-described nucleic acids. The nucleic acid encoding a polypeptide such as an enzyme of the disclosure may be operably linked to a promoter. Specifically where recombinant expression in a filamentous fungal host is desired, the promoter can be a filamentous fungal promoter. The nucleic acids can be, e.g., under the control of heterologous promoters. The nucleic acids can also be expressed under the control of constitutive or inducible promoters. Examples of promoters that can be used include, but are not limited to, a cellulase promoter, a xylanase promoter, the 1818 promoter (previously identified as a highly expressed protein by EST mapping *Trichoderma*). For example, the promoter can suitably be a cellobiohydrolase, endoglucanase, or β -glucosidase promoter. A particularly suitable promoter can be, for example, a *T. reesei* cellobiohydrolase, endoglucanase, or β -glucosidase promoter. For example, the promoter is a cellobiohydrolase I (*cbh1*) promoter. Non-limiting examples of promoters include a *cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl4*, *egl5*, *pki1*, *gpd1*, *xyn1*, or *xyn2* promoter. Additional non-limiting examples of promoters include a *T. reesei* *cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl4*, *egl5*, *pki1*, *gpd1*, *xyn1*, or *xyn2* promoter.

[0204] As used herein, the term “operably linked” means that selected nucleotide sequence (e.g., encoding a polypeptide described herein) is in proximity with a promoter to allow the promoter to regulate expression of the selected DNA. In addition, the promoter is located upstream of the selected nucleotide sequence in terms of the direction of transcription and translation. By “operably linked” is meant that a nucleotide sequence and a regulatory sequence(s) are connected in such a way as to permit gene expression when the appropriate molecules (e.g., transcriptional activator proteins) are bound to the regulatory sequence(s).

[0205] The present disclosure further provides host cells containing any of the polynucleotides vectors, or expression cassettes described herein. The present disclosure also provides host cells that can be used to express one or more polypeptides of the disclosure. Suitable host cells include cells of any microorganism (e.g., cells of a bacterium, a protist, an alga, a fungus (e.g., a yeast or filamentous fungus), or other microbe), and are preferably cells of a bacterium, a yeast, or a filamentous fungus.

[0206] Suitable host cells of the bacterial genera include, but are not limited to, cells of *Escherichia*, *Bacillus*, *Lactobacillus*, *Pseudomonas*, and *Streptomyces*. Suitable cells of bacterial species include, e.g., cells of *Escherichia coli*, *Bacillus subtilis*, *Bacillus licheniformis*, *Lactobacillus brevis*, *Pseudomonas aeruginosa*, or *Streptomyces lividans*.

[0207] Suitable host cells of the genera of yeast include, but are not limited to, cells of *Saccharomyces*, *Schizosaccharomyces*, *Candida*, *Hansenula*, *Pichia*, *Kluyveromyces*, and *Phaffia*. Suitable cells of yeast species include, but are not limited to, cells of *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Candida albicans*, *Hansenula polymorpha*, *Pichia pastoris*, *P. canadensis*, *Kluyveromyces marxianus*, and *Phaffia rhodozyma*.

[0208] Suitable host cells of filamentous fungi include all filamentous forms of the subdivision *Eumycotina*. Suitable cells of filamentous fungal genera include, but are not limited to, cells of *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Corynascus*, *Chaetomium*, *Cryptococcus*, *Filobasidium*, *Fusarium*, *Gibberella*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Mucor*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Scytalidium*, *Schizophyllum*, *Sporotrichum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, and *Trichoderma*. Suitable cells of filamentous fungal species include, but are not limited to, cells of *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*,

Aspergillus oryzae, *Chrysosporium lucknowense*, *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*
5 *sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*,
Fusarium venenatum, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*,
Ceriporiopsis caregiea, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis*
rivulosa, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus*
hirsutus, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora*
10 *thermophila*, *Neurospora crassa*, *Neurospora intermedia*, *Penicillium purpurogenum*,
Penicillium canescens, *Penicillium solitum*, *Penicillium funiculosum* *Phanerochaete*
chrysosporium, *Phlebia radiata*, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*,
Trametes villosa, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*,
Trichoderma longibrachiatum, *Trichoderma reesei*, and *Trichoderma viride*.

[0209] The disclosure provides a host cell, e.g., a recombinant fungal host cell or a recombinant filamentous fungus, engineered to recombinantly express a polypeptide having GH61/endoglucanase activity (e.g., *T. reesei* Eg4 or a variant thereof).

[0210] The present disclosure also provides a recombinant host cell e.g., a recombinant fungal host cell or a recombinant microorganism, e.g., a filamentous fungus, such as a
20 recombinant *T. reesei*, that is engineered to recombinantly express *T. reesei* Xyn3, *T. reesei* Bgl1 (also termed "Tr3A"), Fv3A, Fv43D, and Fv51A polypeptides. For example, the recombinant host cell is suitably a *T. reesei* host cell. The recombinant fungus is suitably a recombinant *T. reesei*. The disclosure provides, for example, a *T. reesei* host cell engineered to recombinantly express *T. reesei* Eg4, *T. reesei* Xyn3, *T. reesei* Bgl1, Fv3A, Fv43D, and
25 Fv51A polypeptides. Alternatively the present disclosure also provides a recombinant host cell or a recombinant microorganism that is, e.g., an *Aspergillus* (such as an *A. oryzae*, *A. niger*) host cell or a recombinant *Aspergillus* engineered to recombinantly express the polypeptides described herein.

[0211] Additionally the disclosure provides a recombinant host cell or recombinant
30 organism that is engineered to express an enzyme blend comprising suitable enzymes in ratios suitable for saccharification. The recombinant host cell is, for example, a fungal host cell or a bacterial host cell. The recombinant fungus is, e.g., a recombinant *T. reesei*, *A.*

oryzae, *A. niger*, or yeast. The recombinant fungal host cell may be, e.g., a *T. reesei*, *A. oryzae*, *A. niger*, or yeast cell. The recombinant bacterial host cell may be, e.g., a *Bacillus subtilis*, or an *E. coli* cell. The recombinant bacterial organism may be, e.g., a *Bacillus subtilis* or an *E. coli*. Examples of enzyme ratios/amounts present in suitable enzyme blends are described herein such as below.

Compositions

[0212] The disclosure also provides compositions (e.g., non-naturally occurring compositions) such as enzyme compositions containing cellulase(s) and/or hemicellulase(s), which can be used to hydrolyze biomass material and/or reduce the viscosity of biomass mixture (e.g., biomass saccharification mixture containing enzyme and substrate).

[0213] Cellulases include enzymes capable of hydrolyzing cellulose (beta-1,4-glucan or beta D-glucosidic linkages) polymers to glucose, cellobiose, cellooligosaccharides, and the like. Cellulases have been traditionally divided into three major classes: endoglucanases (EC 3.2.1.4) ("EG"), exoglucanases or cellobiohydrolases (EC 3.2.1.91) ("CBH") and β -glucosidases (β -D-glucoside glucohydrolase; EC 3.2.1.21) ("BG") (Knowles et al., 1987, Trends in Biotechnology 5(9):255-261; Shulein, 1988, Methods in Enzymology, 160:234-242). Endoglucanases act mainly on the amorphous parts of the cellulose fiber, whereas cellobiohydrolases are also able to degrade crystalline cellulose. Hemicellulases include, for example, xylanases, β -xylosidases, and L- α -arabinofuranosidases.

[0214] The composition of the invention may be a multi-enzyme blend, comprising more than one enzyme. The enzyme composition of the invention can suitably include one or more additional enzymes derived from other microorganisms, plants, or organisms. Synergistic enzyme combinations and related methods are contemplated. The disclosure includes methods for identifying the optimum ratios of the enzymes included in the enzyme compositions for degrading various types of biomass materials. These methods include, e.g., tests to identify the optimum proportion or relative weights of enzymes to be included in the enzyme composition of the invention in order to effectuate efficient conversion of various substrates (e.g., lignocellulosic substrates) to their constituent fermentable sugars.

[0215] The cell walls of higher plants are comprised of a variety of carbohydrate polymer (CP) components. These CP interact through covalent and non-covalent means, providing the structural integrity required to form rigid cell walls and resist turgor pressure in plants. The major CP found in plants is cellulose, which forms the structural backbone of the cell

wall. During cellulose biosynthesis, chains of poly- β -1,4-D-glucose self associate through hydrogen bonding and hydrophobic interactions to form cellulose microfibrils, which further self-associate to form larger fibrils. Cellulose microfibrils are often irregular structurally and contain regions of varying crystallinity. The degree of crystallinity of cellulose fibrils
5 depends on how tightly ordered the hydrogen bonding is between and among its component cellulose chains. Areas with less-ordered bonding, and therefore more accessible glucose chains, are referred to as amorphous regions. The general model for cellulose depolymerization to glucose involves a minimum of three distinct enzymatic activities.

Endoglucanases cleave cellulose chains internally to shorter chains in a process that increases
10 the number of accessible ends, which are more susceptible to exoglucanase activity than the intact cellulose chains. These exoglucanases (*e.g.*, cellobiohydrolases) are specific for either reducing ends or non-reducing ends, liberating, in most cases, cellobiose, the dimer of glucose. The accumulating cellobiose is then subject to cleavage by cellobiases (*e.g.*, β -1,4-glucosidases) to glucose. Cellulose contains only anhydro-glucose. In contrast,

15 hemicellulose contains a number of different sugar monomers. For instance, aside from glucose, sugar monomers in hemicellulose can also include xylose, mannose, galactose, rhamnose, and arabinose. Hemicelluloses mostly contain D-pentose sugars and occasionally small amounts of L-sugars. Xylose is typically present in the largest amount, but mannuronic acid and galacturonic acid also tend to be present. Hemicelluloses include xylan,
20 glucuronoxylan, arabinoxylan, glucomannan, and xyloglucan.

[0216] The compositions (*e.g.*, enzymes and multi-enzyme compositions) of the disclosure can be used for saccharification of cellulose materials (*e.g.*, glucan) and/or hemicellulose materials (*e.g.*, xylan, arabinoxylan, and xylan- or arabinoxylan-containing substrates). The enzyme blend/composition is suitably a non-naturally occurring composition.

25 **[0217]** The enzyme compositions provided herein may comprise a mixture of xylan-hydrolyzing, hemicellulose- and/or cellulose-hydrolyzing enzymes, which include at least one, several, or all of a cellulase, including a glucanase; a cellobiohydrolase; an L- α -arabinofuranosidase; a xylanase; a β -glucosidase; and a β -xylosidase. The present disclosure also provides enzyme compositions that may be non-naturally occurring compositions. As
30 used herein, the term “enzyme compositions” refers to: (1) a composition made by combining component enzymes, whether in the form of a fermentation broth or partially or completely isolated or purified; (2) a composition produced by an organism modified to express one or

more component enzymes; in certain embodiments, the organism used to express one or more component enzymes can be modified to delete one or more genes; in certain other embodiments, the organism used to express one or more component enzymes can further comprise proteins affecting xylan hydrolysis, hemicellulose hydrolysis, and/or cellulose hydrolysis; (3) a composition made by combining component enzymes simultaneously, separately, or sequentially during a saccharification or fermentation reaction; (4) an enzyme mixture produced in situ, *e.g.*, during a saccharification or fermentation reaction; (5) a composition produced in accordance with any or all of the above (1)-(4).

[0218] The term “fermentation broth” as used herein refers to an enzyme preparation produced by fermentation that undergoes no or minimal recovery and/or purification subsequent to fermentation. For example, microbial cultures are grown to saturation, incubated under carbon-limiting conditions to allow protein synthesis (*e.g.*, expression of enzymes). Then, once the enzyme(s) are secreted into the cell culture media, the fermentation broths can be used. The fermentation broths of the disclosure can contain unfractionated or fractionated contents of the fermentation materials derived at the end of the fermentation. For example, the fermentation broths of the invention are unfractionated and comprise the spent culture medium and cell debris present after the microbial cells (*e.g.*, filamentous fungal cells) undergo a fermentation process. The fermentation broth can suitably contain the spent cell culture media, extracellular enzymes, and live or killed microbial cells. Alternatively, the fermentation broths can be fractionated to remove the microbial cells. In those cases, the fermentation broths can, for example, comprise the spent cell culture media and the extracellular enzymes.

[0219] The enzyme compositions such as cellulase compositions provided herein may be capable of achieving at least 0.1 (*e.g.* 0.1 to 0.4) fraction product as determined by the calcofluor assay. All chemicals used were of analytical grade. Avicel PH-101 was purchased from FMC BioPolymer (Philadelphia, PA). Cellobiose and calcofluor white were purchased from Sigma (St. Louise, MO). Phosphoric acid swollen cellulose (PASC) was prepared from Avicel PH-101 using an adapted protocol of Walseth, TAPPI 1971, 35:228 and Wood, Biochem. J. 1971, 121:353-362. In short, Avicel was solubilized in concentrated phosphoric acid then precipitated using cold deionized water. After the cellulose is collected and washed with more water to neutralize the pH, it was diluted to 1% solids in 50 mM sodium acetate pH5. All enzyme dilutions were made into 50 mM sodium acetate buffer, pH5.0. GC220

Cellulase (Danisco US Inc., Genencor) was diluted to 2.5, 5, 10, and 15 mg protein/G PASC, to produce a linear calibration curve. Samples to be tested were diluted to fall within the range of the calibration curve, *i.e.* to obtain a response of 0.1 to 0.4 fraction product. 150 μ L of cold 1% PASC was added to 20 μ L of enzyme solution in 96-well microtiter plates. The plate was covered and incubated for 2 h at 50°C, 200 rpm in an Innova incubator/shaker. The reaction was quenched with 100 μ L of 50 μ g/mL Calcofluor in 100 mM Glycine, pH10. Fluorescence was read on a fluorescence microplate reader (SpectraMax M5 by Molecular Devices) at excitation wavelength Ex = 365 nm and emission wavelength Em = 435 nm. The result is expressed as the fraction product according to the equation:

$$FP = 1 - (FI_{\text{sample}} - FI_{\text{buffer w/ cellobiose}}) / (FI_{\text{zero enzyme}} - FI_{\text{buffer w/ cellobiose}}),$$

wherein FP is fraction product, and FI = fluorescence units.

[0220] Any of the enzymes described specifically herein can be combined with any one or more of the enzymes described herein or with any other available and suitable enzymes, to produce a suitable multi-enzyme blend/composition. The disclosure is not restricted or limited to the specific exemplary combinations listed below.

Exemplary compositions

[0221] There are provided non-naturally occurring compositions comprising a polypeptide having GH61/endoglucanase activity. The invention also provides a non-naturally occurring composition comprising whole cellulase comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, whole cellulase enriched with a polypeptide having GH61/endoglucanase activity such as endoglucanase IV (*e.g.*, *T. reesei* Eg4 polypeptide-enriched whole cellulase)). The polypeptide having GH61/endoglucanase activity may be any polypeptide having GH61/endoglucanase activity provided herein. In some aspects, the polypeptide having GH61/endoglucanase activity is *T. reesei* Eg4 or a variant thereof. A variant of *T. reesei* Eg4 can be any of the variants provided above.

[0222] Endoglucanase is referred to herein as “Eg” or “Egl,” interchangeably, in the present disclosure including figures.

[0223] As used herein, the term “naturally occurring composition” refers to a composition produced by a naturally occurring source, comprising one or more enzymatic components or activities, wherein each of the components or activities is found at the ratio and level produced by the naturally-occurring source as it is found in nature, untouched, unmodified by

the human hand. Accordingly, a naturally occurring composition is, e.g., one that is produced by an organism unmodified with respect to the cellulolytic or hemicellulolytic enzymes such that the ratio or levels of the component enzymes are unaltered from that produced by the native organism in its native environment. A “non-naturally occurring composition,” on the other hand, refers to a composition produced by: (1) combining component cellulolytic or hemicellulolytic enzymes either in a naturally occurring ratio or a non-naturally occurring, *i.e.*, altered, ratio; or (2) modifying an organism to express, overexpress or underexpress one or more endogenous or exogenous enzymes; or (3) modifying an organism such that at least one endogenous enzyme is deleted. A “non-naturally occurring composition” also refers to a composition produced by a naturally-occurring, unmodified organism, but cultured in a man-made medium or environment that is different from the organism’s native environment such that the amounts of enzymes in the composition differ from those existing in a composition made by a native organism grown in its native habitat.

[0224] Any one of GH61 endoglucanase polypeptides or a variant thereof may be used in any of the compositions described herein. A suitable GH61 endoglucanase may include one of the polypeptides shown in FIG.1 of the present disclosure. Suitable GH61 endoglucanases include those that are represented by their GenBank Accession Numbers CAB97283.2, CAD70347.1, CAD21296.1, CAE81966.1, CAF05857.1, EAA26873.1, EAA29132.1, EAA30263.1, EAA33178.1, EAA33408.1, EAA34466.1, EAA36362.1, EAA29018.1, and EAA29347.1, or St61 from *S.thermophilum* 24630, St61A from *S.thermophilum* 23839c, St61B from *S.thermophilum* 46583, St61D from *S.thermophilum* 80312, Afu61a from *A. fumigatus* Afu3g03870 (NCBI Ref: XP_748707), an endoglucanase of NCBI Ref: XP_750843.1 from *A.fumigatus* Afu6g09540, an endoglucanase of *A. fumigatus* EDP47167, an endoglucanase of *T. terrestris* 16380, an endoglucanase of *T.terrestris* 155418, an endoglucanase of *T.terrestris* 68900, Cg61A (EAQ86340.1) from *C.globosum*, *T. reesei* Eg7, *T. reesei* Eg4, and an endoglucanase with GenBank Accession: XP_752040 from *A.fumigatus* Af293. In some aspects, the polypeptide having GH61/ endoglucanase activity (*e.g.*, isolated polypeptide) is a variant of GH61 endoglucanase or EG IV.

[0225] In some aspects, the polypeptide having GH61/endoglucanase activity (including a variant of GH61 endoglucanase) is one comprising any one of SEQ ID NOs: 1-29 and 148, or one that comprises a polypeptide having at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 92.5%, 95%, 96%, 97%, 98%, or 99% sequence identity to any one of SEQ ID NOs: 1-

29 and 148. In some aspects, the polypeptide having GH61/endoglucanase activity (including a variant of GH61 endoglucanase) may comprise at least one motif (at least any of 2, 3, 4, 5, 6, 7, or 8) selected from SEQ ID NOs:84-91. It may comprise one or more sequence motif(s) selected from the group consisting of: (1) SEQ ID NOs:84 and 88; (2) SEQ ID NOs:85 and 88; (3) SEQ ID NO:86; (4) SEQ ID NO:87; (5) SEQ ID NOs:84, 88 and 89; (6) SEQ ID NOs:85, 88, and 89; (7) SEQ ID NOs: 84, 88, and 90; (8) SEQ ID NOs: 85, 88 and 90; (9) SEQ ID NOs:84, 88 and 91; (10) SEQ ID NOs: 85, 88 and 91; (11) SEQ ID NOs: 84, 88, 89 and 91; (12) SEQ ID NOs: 84, 88, 90 and 91; (13) SEQ ID NOs: 85, 88, 89 and 91; and (14) SEQ ID NOs: 85, 88, 90 and 91 .

[0226] In some aspects of any one of the compositions or methods described herein, the polypeptide having GH61/endoglucanase activity (including a variant of GH61 endoglucanase) may have at least about 60% (*e.g.*, at least about any of 60%, 65%, 70%, 75%, 80%, 85%, 88%, 90%, 92.5%, 95%, 96%, 97%, 98%, or 99%) sequence identity to residues 22 to 344 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to at least about 5 residues (*e.g.*, at least about any of 6, 7, 8, 9, 10, 11, or 12) of H22, D61, G63, C77, H107, R177, E179, H184, Q193, C198, Y195, and Y232 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to H22, D61, G63, C77, H107, R177, E179, H184, Q193, C198, Y195, and Y232 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to at least 5 residues (*e.g.*, at least about any of 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, or 19) of G313, Q314, C315, G316, G317, S321, G322, P323, T324, C326, A327, T331, C332, N336, Y338, Y339, Q341, C342, and L343 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises residues corresponding to G313, Q314, C315, G316, G317, S321, G322, P323, T324, C326, A327, T331, C332, N336, Y338, Y339, Q341, C342, and L343 of SEQ ID NO:27. In some aspects, the polypeptide or a variant thereof comprises a CBM domain (*e.g.*, functional CBM domain). In some aspects, the polypeptide or a variant thereof comprises a catalytic domain (*e.g.*, functional catalytic domain). In some aspects, the polypeptide or a variant thereof is isolated. In some aspects, the polypeptide or a variant thereof has endoglucanase activity.

[0227] In some aspects, the polypeptide having GH61/endoglucanase activity is endoglucanase IV, for example, a *T. reesei* Eg4 polypeptide or a variant thereof. For example, the disclosure provides non-naturally occurring compositions comprising a *T. reesei*

Eg4 polypeptide or a variant thereof. A variant of *T. reesei* Eg4 polypeptide can be any one of the variants of *T. reesei* Eg4 polypeptide described herein. In some aspects, the polypeptide having GH61/endoglucanase activity includes amino acid sequence SEQ ID NO:27 or residues 22 to 344 of SEQ ID NO:27.

5 [0228] In some aspects, there is provided a composition comprising an isolated (or substantially purified) polypeptide having glycosyl hydrolase family 61 (“GH61”)/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof). Methods of producing polypeptide, recovering the polypeptide, and isolating or purifying the polypeptide are known to one of skill in the art.

10 [0229] In some aspects of any of the compositions or methods described herein, the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is expressed from a host cell, wherein the nucleic acid encoding the polypeptide having GH61/endoglucanase activity has been engineered into the host cell. In some aspects, the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is
15 heterologous to the host cell expressing the polypeptide having GH61/endoglucanase activity.

[0230] The present disclosure provides compositions comprising a polypeptide having GH61/endoglucanase activity and comprising at least one cellulase polypeptide and/or at least one hemicellulase polypeptide, or a mixture thereof. In some aspects, the composition comprises at least one (*e.g.*, at least 2, 3, 4, 5, 6, 7, or 8) cellulase polypeptide(s). In some
20 aspects, the cellulase polypeptide is a polypeptide having endoglucanase activity, a polypeptide having cellobiohydrolase activity, or a polypeptide having β -glucosidase activity. In some aspects, the composition comprises at least one (*e.g.*, at least 2, 3, 4, 5, 6, 7, or 8) hemicellulase polypeptide(s). In some aspects, the hemicellulase polypeptide is a polypeptide having xylanase activity, a polypeptide having β -xylosidase activity, or a polypeptide having
25 L- α -arabinofuranosidase activity. In some aspects, the composition further comprises at least one (*e.g.*, at least 2, 3, 4, 5, 6, 7, or 8) cellulase polypeptide(s) and at least one (*e.g.*, at least 2, 3, 4, 5, 6, 7, or 8) hemicellulase polypeptide(s). Varying amounts for polypeptide(s) included in the compositions provided herein are provided below in “Amount of component(s) in compositions” section.

30 [0231] Cellulases and hemicellulases for use in accordance with the methods and compositions of the disclosure can be obtained from, or produced recombinantly from, *inter alia*, one or more of the following organisms: *Crinipellis scapella*, *Macrophomina*

phaseolina, *Myceliophthora thermophila*, *Sordaria fimicola*, *Volutella colletotrichoides*, *Thielavia terrestris*, *Acremonium* sp., *Exidia glandulosa*, *Fomes fomentarius*, *Spongipellis* sp., *Rhizophlyctis rosea*, *Rhizomucor pusillus*, *Phycomyces niteus*, *Chaetostylum fresenii*, *Diplodia gossypina*, *Ulospora bilgramii*, *Saccobolus dilutellus*, *Penicillium verruculosum*,
 5 *Penicillium chrysogenum*, *Thermomyces verrucosus*, *Diaporthe syngenesia*, *Colletotrichum lagenarium*, *Nigrospora* sp., *Xylaria hypoxylon*, *Nectria pinea*, *Sordaria macrospora*, *Thielavia thermophila*, *Chaetomium mororum*, *Chaetomium virscens*, *Chaetomium brasiliensis*, *Chaetomium cunicolorum*, *Syspastospora boninensis*, *Cladorrhinum foecundissimum*, *Scytalidium thermophila*, *Gliocladium catenulatum*, *Fusarium oxysporum*
 10 ssp. *lycopersici*, *Fusarium oxysporum* ssp. *passiflora*, *Fusarium solani*, *Fusarium anguioides*, *Fusarium poae*, *Humicola nigrescens*, *Humicola grisea*, *Panaeolus retirugis*, *Trametes sanguinea*, *Schizophyllum commune*, *Trichothecium roseum*, *Microsphaeropsis* sp., *Acsobolus stictoideus* spej., *Poronia punctata*, *Nodulisporum* sp., *Trichoderma* sp. (e.g., *Trichoderma reesei*) and *Cylindrocarpon* sp.

15 **[0232]** In the present disclosure, the cellulase or hemicellulase may be prepared from any known microorganism cultivation method(s), resulting in the expression of enzymes capable of hydrolyzing a cellulosic material. Fermentation may include shake flask cultivation, small- or large-scale fermentation, such as continuous, batch, fed-batch, or solid state fermentations in laboratory or industrial fermenters performed in a suitable medium and under conditions
 20 allowing the cellulase to be expressed or isolated. Generally, the microorganism is cultivated in a cell culture medium suitable for production of enzymes capable of hydrolyzing a cellulosic material. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable culture media, temperature ranges and other conditions suitable for growth and cellulase
 25 production are known in the art. As a non-limiting example, the normal temperature range for the production of cellulases by *T. reesei* is 24°C to 28°C.

[0233] The present disclosure provides non-naturally occurring compositions comprising a polypeptide having GH61/endoglucanase activity (e.g., endoglucanase IV polypeptide such as *T. reesei* Eg4 polypeptide or a variant thereof), wherein the composition further comprises
 30 at least 1 polypeptide having endoglucanase activity (e.g., at least 2, 3, 4, or 5 polypeptides having endoglucanase activity), at least 1 polypeptide having cellobiohydrolase activity (e.g., at least 2, 3, 4, or 5 polypeptides having cellobiohydrolase activity), at least 1 polypeptide

having glucosidase activity (*e.g.*, β -glucosidase) (*e.g.*, at least 2, 3, 4, or 5 polypeptides having β -glucosidase activity), at least 1 polypeptide having xylanase activity (*e.g.*, at least 2, 3, 4, or 5 polypeptides having xylanase activity), at least 1 polypeptide having xylosidase activity (*e.g.*, β -xylosidase) (*e.g.*, at least 2, 3, 4, or 5 polypeptides having β -xylosidase activity), and/or at least 1 polypeptide having arabinofuranosidase activity (*e.g.*, L- α -arabinofuranosidase) (*e.g.*, at least 2, 3, 4, or 5 polypeptides having L- α -arabinofuranosidase activity). Varying amounts for polypeptide(s) included in the compositions provided herein are provided below in “Amount of component(s) in compositions” section.

[0234] The present disclosure provides non-naturally occurring compositions comprising whole cellulase comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, whole cellulase enriched with endoglucanase IV polypeptide, such as, *e.g.*, *T. reesei* Eg4 polypeptide or a variant thereof), wherein the composition further comprises at least 1 polypeptide having endoglucanase activity (*e.g.*, at least 2, 3, 4, or 5 polypeptides having endoglucanase activity), at least 1 polypeptide having cellobiohydrolase activity (*e.g.*, at least 2, 3, 4, or 5 polypeptides having cellobiohydrolase activity), at least 1 polypeptide having glucosidase activity (*e.g.*, β -glucosidase) (*e.g.*, at least 2, 3, 4, or 5 polypeptides having β -glucosidase activity), at least 1 polypeptide having xylanase activity (*e.g.*, at least 2, 3, 4, or 5 polypeptides having xylanase activity), at least one polypeptide having xylosidase activity (*e.g.*, β -xylosidase) (*e.g.*, at least 2, 3, 4, or 5 polypeptides having β -xylosidase activity), and/or at least one polypeptide having arabinofuranosidase activity (*e.g.*, L- α -arabinofuranosidase) (*e.g.*, at least 2, 3, 4, or 5 polypeptides having L- α -arabinofuranosidase activity). Varying amounts for polypeptide(s) included in the compositions provided herein are provided below in “Amount of component(s) in compositions” section.

[0235] In some aspects, the composition comprises a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least 1 polypeptide having xylanase activity (*e.g.*, *T. reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, AfuXyn5, or a variant thereof). In some aspects, the polypeptide having xylanase activity is *T. reesei* Xyn3. The composition may further comprise at least 1 polypeptide having β -glucosidase activity (*e.g.*, Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, and/or Tn3B). The composition may further comprise at least 1 polypeptide having β -glucosidase activity (*e.g.*, Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B, and/or a variant thereof). The composition may further

comprise at least 1 polypeptide having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof). The composition may further comprise at least 1 polypeptide having endoglucanase activity (*e.g.*, *T. reesei* EG1 (or a variant thereof) and/or *T. reesei* EG2 (or a variant thereof)).

[0236] In some aspects, the composition comprises a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least 1 polypeptide having β -glucosidase activity (*e.g.*, Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B, or a variant thereof). The composition may comprise a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least 1 polypeptide (or at least 2 polypeptides) having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof). The composition may comprise a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and further comprises at least 1 polypeptide (or at least 2 polypeptides) having endoglucanase activity (*e.g.*, *T. reesei* EG1 (or a variant thereof) and/or *T. reesei* EG2 (or a variant thereof)). The composition may comprise a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least 1 polypeptide (or at least two polypeptides) having β -xylosidase activity (*e.g.*, Fv3A, Fv43A, Pf43A, Fv43D, Fv39A, Fv43E, Fo43A, Fv43B, Pa51A, Gz43A, and/or *T. reesei* Bx11). The composition may comprise a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least 1 polypeptide (or at least 2 polypeptides) having β -xylosidase activity (*e.g.*, Fv3A, Fv43A, Pf43A, Fv43D, Fv39A, Fv43E, Fo43A, Fv43B, Pa51A, Gz43A, *T. reesei* Bx11, and/or a variant thereof). The composition may comprise a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one polypeptide (at least 2 polypeptides) having L- α -arabinofuranosidase activity (*e.g.*, Af43A, Fv43B, Pf51A, Pa51A, Fv51A, or a variant thereof).

[0237] In some aspects, any of the polypeptides described herein (*e.g.*, polypeptide having endoglucanase activity, polypeptide having cellobiohydrolase activity, polypeptide having glucosidase activity (*e.g.*, β -glucosidase), polypeptide having xylanase activity, polypeptide having xylosidase activity (*e.g.*, β -xylosidase), or polypeptide having arabinofuranosidase activity (*e.g.*, L- α -arabinofuranosidase)) may be a component of a whole cellulase such as a

whole cellulase described herein. Any of the polypeptides described herein may be produced by expressing an endogenous or exogenous gene encoding the corresponding polypeptide(s). The polypeptide(s) can be, in some circumstances, overexpressed or underexpressed.

[0238] Regarding any of the compositions described above, varying amounts for

5 polypeptide(s) included in the compositions are provided below in “Amount of component(s) in compositions” section.

Polypeptide having endoglucanase activity

[0239] A polypeptide having endoglucanase activity includes a polypeptide that catalyzes the cleavage of internal β -1,4 linkages. Endoglucanase (“EG”) refers to a group of cellulase
10 enzymes classified as EC 3.2.1.4. An EG enzyme hydrolyzes internal beta-1,4 glucosidic bonds of the cellulose. EG catalyzes endohydrolysis of 1,4-beta-D-glycosidic linkages in cellulose, cellulose derivatives (for example, carboxy methyl 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. EG activity can be determined using
15 carboxymethyl cellulose (CMC) hydrolysis according to the procedure of Ghose, 1987, Pure and Appl. Chem. 59: 257-268. In some aspects, at least one polypeptide having endoglucanase activity includes *T. reesei* EG1 (GenBank Accession No. HM641862.1) and/or *T. reesei* EG2 polypeptide (GenBank Accession No. ABA64553.1).

[0240] A thermostable *T. terrestris* endoglucanase (Kvesitadaze *et al.*, Applied Biochem.
20 Biotech. 1995, 50:137-143) is, in another example, used in the methods and compositions of the present disclosure. Moreover, a *T. reesei* EG3 (GenBank Accession No. AAA34213.1) (Okada *et al.* Appl. Environ. Microbiol. 1988, 64:555-563), EG5 (GenBank Accession No. AAP57754) (Saloheimo *et al.* Molecular Microbiology 1994, 13:219-228), EG6 (FIG. 89A) (U.S. Patent Publication No. 20070213249), or EG7 (GenBank Accession No. AAP57753)
25 (U.S. Patent Publication No. 20090170181), an *A. cellulolyticus* EI endoglucanase (Swiss-Prot entry P54583.1) (U.S. Pat. No. 5,536,655), a *H. insolens* endoglucanase V (EGV) (Protein Data Bank entry 4ENG), a *S. coccosporum* endoglucanase (FIG. 89B) (U.S. Patent Publication No. 20070111278), an *A. aculeatus* endoglucanase F1-CMC (Swiss-Prot entry P22669.1) (Ooi *et al.* Nucleic Acid Res. 1990, 18:5884), an *A. kawachii* IFO 4308
30 endoglucanase CMCase-1 (Swiss-Prot entry Q96WQ8.1) (Sakamoto *et al.* Curr. Genet. 1995, 27:435-439), an *E. carotovora* endoglucanase CelS (GenBank Accession No. AAA24817.1) (Saarilahti *et al.* Gene 1990, 90:9-14); or an *A. thermophilum* ALKO4245 endoglucanase

(U.S. Patent Publication No. 20070148732) can also be used. Additional suitable endoglucanases are described in, *e.g.*, WO 91/17243, WO 91/17244, WO 91/10732, U.S. Patent No. 6,001,639. A polypeptide having endoglucanase activity may be a variant of any one of the endoglucanases provided herein.

5 *Polypeptide having cellobiohydrolase activity*

[0241] A polypeptide having cellobiohydrolase activity includes a polypeptide having 1,4-D-glucan cellobiohydrolase (E.C. 3.2.1.91) activity which catalyzes the hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, cellotetraose, or any beta-1,4-linked glucose containing polymer, releasing cellobiose from the ends of the chain. For purposes of the present invention, cellobiohydrolase activity can be determined by release of water-soluble reducing sugar from cellulose as measured by the PHBAH method of Lever et al., 1972, Anal. Biochem. 47: 273-279. A distinction between the exoglucanase mode of attack of a cellobiohydrolase and the endoglucanase mode of attack can be made by a similar measurement of reducing sugar release from substituted cellulose such as carboxymethyl cellulose or hydroxyethyl cellulose (Ghose, 1987, Pure & Appl. Chem. 59: 257-268). A true cellobiohydrolase will have a very high ratio of activity on unsubstituted versus substituted cellulose (Bailey et al, 1993, Biotechnol. Appl. Biochem. 17: 65-76).

[0242] Suitable CBHs can be selected from *A. bisporus* CBH1 (Swiss Prot Accession No. Q92400), *A. aculeatus* CBH1 (Swiss Prot Accession No. O59843), *A. nidulans* CBHA (GenBank Accession No. AF420019) or CBHB (GenBank Accession No. AF420020), *A. niger* CBHA (GenBank Accession No. AF156268) or CBHB (GenBank Accession No. AF156269), *C. purpurea* CBH1 (Swiss Prot Accession No. O00082), *C. carbonarum* CBH1 (Swiss Prot Accession No. Q00328), *C. parasitica* CBH1 (Swiss Prot Accession No. Q00548), *F. oxysporum* CBH1 (Cel7A) (Swiss Prot Accession No. P46238), *H. grisea* CBH1.2 (GenBank Accession No. U50594), *H. grisea* var. *thermoidea* CBH1 (GenBank Accession No. D63515), CBH1.2 (GenBank Accession No. AF123441), or exo1 (GenBank Accession No. AB003105), *M. albomyces* Cel7B (GenBank Accession No. AJ515705), *N. crassa* CBHI (GenBank Accession No. X77778), *P. funiculosus* CBHI (Cel7A) (GenBank Accession No. AJ312295) (U.S. Patent Publication No. 20070148730), *P. janthinellum* CBHI (GenBank Accession No. S56178), *P. chrysosporium* CBH (GenBank Accession No. M22220), or CBHI-2 (Cel7D) (GenBank Accession No. L22656), *T. emersonii* CBH1A (GenBank Accession No. AF439935), *T. viride* CBH1 (GenBank Accession No. X53931), or

V. volvacea V14 CBH1 (GenBank Accession No. AF156693). A polypeptide having cellobiohydrolase activity may be a variant of any one of CBHs provided herein.

[0243] In some aspects, at least one polypeptide having cellobiohydrolase activity includes *T.reesei* CBH 1 (Swiss-Prot entry P62694.1) (or a variant thereof) and/or *T. reesei* CBH2

5 (Swiss-Prot entry P07987.1) (or a variant thereof) polypeptide. See Shoemaker et al. Bio/Technology 1983, 1:691-696; see also Teeri et al. Bio/Technology 1983, 1:696-699, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, which are *T. reesei* CBH1 homologs; *T. terrestris* 6A, *S. thermophile* 6A, 6B, which are *T. reesei* CBH2 homologs, or a variant thereof.

10 *Polypeptide having glucosidase activity*

[0244] A polypeptide having glucosidase activity includes a polypeptide having beta-D-glucoside glucohydrolase (E.C. 3.2.1.21) activity which catalyzes the hydrolysis of cellobiose with the release of beta-D-glucose. For purposes of the present invention, β -glucosidase activity may be measured by methods known in the art, e.g., HPLC. A polypeptide having

15 glucosidase activity includes members of certain GH families, including, without limitation, members of GH families 1, 3, 9 or 48, which catalyze the hydrolysis of cellobiose to release β -D-glucose. A polypeptide having glucosidase activity includes β -glucosidase such as β -glucosidase obtained from a number of microorganisms, by recombinant means, or be

20 purchased from commercial sources. Examples of β -glucosidases from microorganisms include, without limitation, ones from bacteria and fungi. For example, a β -glucosidase is suitably obtained from a filamentous fungus. In some aspects, at least one polypeptide having glucosidase activity (e.g., β -glucosidase activity) is a *T. reesei* Bgl1 polypeptide.

[0245] The β -glucosidases can be obtained, or produced recombinantly, from, *inter alia*, *A. aculeatus* (Kawaguchi et al. Gene 1996, 173: 287-288), *A. kawachi* (Iwashita et al. Appl.

25 Environ. Microbiol. 1999, 65: 5546-5553), *A. oryzae* (WO 2002/095014), *C. biazotea* (Wong et al. Gene, 1998, 207:79-86), *P. funiculosum* (WO 2004/078919), *S. fibuligera* (Machida et al. Appl. Environ. Microbiol. 1988, 54: 3147-3155), *S. pombe* (Wood et al. Nature 2002, 415: 871-880), or *T. reesei* (e.g., β -glucosidase 1 (U.S. Patent No. 6,022,725), β -glucosidase 3 (U.S. Patent No.6,982,159), β - glucosidase 4 (U.S. Patent No. 7,045,332), β -glucosidase 5

30 (US Patent No. 7,005,289), β -glucosidase 6 (U.S. Publication No. 20060258554), β -glucosidase 7 (U.S. Publication No. 20060258554)). A polypeptide having β -glucosidases activity may be a variant of any one of β -glucosidases provided herein.

[0246] The β -glucosidase can be produced by expressing an endogenous or exogenous gene encoding a β -glucosidase. For example, β -glucosidase can be secreted into the extracellular space *e.g.*, by Gram-positive organisms (*e.g.*, *Bacillus* or *Actinomycetes*), or a eukaryotic hosts (*e.g.*, *Trichoderma*, *Aspergillus*, *Saccharomyces*, or *Pichia*). The β -glucosidase can be, in some circumstances, overexpressed or underexpressed.

[0247] The β -glucosidase can also be obtained from commercial sources. Examples of commercial β -glucosidase preparation suitable for use include, *e.g.*, *T. reesei* β -glucosidase in Accellerase[®] BG (Danisco US Inc., Genencor); NOVOZYM[™] 188 (a β -glucosidase from *A. niger*); *Agrobacterium sp.* β -glucosidase, and *T. maritima* β -glucosidase from Megazyme (Megazyme International Ireland Ltd., Ireland.).

[0248] β -glucosidase activity can be determined by a number of suitable means known in the art, such as the assay described by Chen *et al.*, in Biochimica et Biophysica Acta 1992, 121:54-60, wherein 1 pNPG denotes 1 μ mol of Nitrophenol liberated from 4-nitrophenyl- β -D-glucopyranoside in 10 min at 50°C (122°F) and pH 4.8.

Polypeptide having xylanase activity

[0249] Xylanase activity may be measured by using colorimetric azo-birchwood xylan assay (S-AXBL, Megazyme International Ireland Ltd., Ireland).

[0250] A polypeptide having xylanase activity may include Group A xylanases, selected from, *e.g.*, Xyn, Xyn2, AfuXyn2, and/or AfuXyn5 polypeptide, or a variant thereof.

[0251] Any of the compositions described herein may optionally comprise one or more xylanases in addition to or in place of the one or more Group A xylanases. Any xylanase (EC 3.2.1.8) can be used as the additional one or more xylanases. Suitable xylanases include, *e.g.*, *C. saccharolyticum* xylanase (Luthi *et al.* 1990, Appl. Environ. Microbiol. 56(9):2677-2683), *T. maritima* xylanase (Winterhalter & Liebel, 1995, Appl. Environ. Microbiol. 61(5):1810–1815), *Thermatoga Sp.* Strain FJSS-B.1 xylanase (Simpson *et al.* 1991, Biochem. J. 277, 413–417), *B. circulans* xylanase (BcX) (U.S. Patent No. 5,405,769), *A. niger* xylanase (Kinoshita *et al.* 1995, Journal of Fermentation and Bioengineering 79(5):422-428), *S. lividans* xylanase (Shareck *et al.* 1991, Gene 107:75-82; Morosoli *et al.* 1986 Biochem. J. 239:587-592; Kluepfel *et al.* 1990, Biochem. J. 287:45-50), *B. subtilis* xylanase (Bernier *et al.* 1983, Gene 26(1):59–65), *C. fimi* xylanase (Clarke *et al.*, 1996, FEMS Microbiology Letters 139:27-35), *P. fluorescens* xylanase (Gilbert *et al.* 1988, Journal of General Microbiology 134:3239–3247), *C. thermocellum* xylanase (Dominguez *et al.*, 1995, Nature Structural Biology 2:569–

576), *B.pumilus* xylanase (Nuyens *et al.* Applied Microbiology and Biotechnology 2001, 56:431-434; Yang *et al.* 1998, Nucleic Acids Res. 16(14B):7187), *C.acetobutylicum* P262 xylanase (Zappe *et al.* 1990, Nucleic Acids Res. 18(8):2179), or *T.harzianum* xylanase (Rose *et al.* 1987, J. Mol. Biol.194(4):755-756). A polypeptide having xylanase activity may be a
 5 variant of any one of the xylanases provided herein.

Polypeptide having xylosidase (e.g., β -xylosidase) activity

[0252] Xylosidase (e.g., β -xylosidase) activity may be measured by using chromogenic substrate 4-nitrophenyl beta-D-xylopyranoside (pNPX, Sigma-Aldrich N2132).

[0253] A polypeptide having xylosidase (e.g., β -xylosidase) activity may be a Group 1 β -
 10 xylosidase enzyme (e.g., Fv3A or Fv43A) or a Group 2 β -xylosidase enzyme (e.g., Pf43A, Fv43D, Fv39A, Fv43E, Fo43A, Fv43B, Pa51A, Gz43A, *T. reesei* Bx11, or a variant thereof). In some aspects, any of the composition provided herein may suitably comprise one or more Group 1 β -xylosidases and one or more Group 2 β -xylosidases.

[0254] Any of the composition provided herein such as the enzyme blends/compositions of
 15 the disclosure can optionally comprise one or more β -xylosidases, in addition to or in place of the Group 1 and/or Group 2 β -xylosidases above. Any β -xylosidase (EC 3.2.1.37) can be used as the additional β -xylosidases. Suitable β -xylosidases include, for example, *T.emersonii* Bx11 (Reen *et al.* 2003, Biochem Biophys Res Commun. 305(3):579-85), *G. stearothermophilus* β -xylosidases (Shallom *et al.* 2005, Biochemistry 44:387-397), *S.*
 20 *thermophilum* β -xylosidases (Zanoelo *et al.* 2004, J. Ind. Microbiol. Biotechnol. 31:170-176), *T.lignorum* β -xylosidases (Schmidt, 1998, Methods Enzymol. 160:662-671), *A. awamori* β -xylosidases (Kurakake *et al.* 2005, Biochim. Biophys. Acta 1726:272-279), *A. versicolor* β -xylosidases (Andrade *et al.* 2004, Process Biochem. 39:1931-1938), *Streptomyces sp.* β -xylosidases (Pinphanichakarn *et al.* 2004, World J. Microbiol. Biotechnol. 20:727-733), *T.*
 25 *maritima* β -xylosidases (Xue and Shao, 2004, Biotechnol. Lett. 26:1511-1515), *Trichoderma sp.* SY β -xylosidases (Kim *et al.* 2004, J. Microbiol. Biotechnol. 14:643-645), *A. niger* β -xylosidases (Oguntimein and Reilly, 1980, Biotechnol. Bioeng. 22:1143-1154), or *P.wortmanni* β -xylosidases (Matsuo *et al.* 1987, Agric. Biol. Chem. 51:2367-2379). A polypeptide having xylosidase (e.g., β -xylosidase) activity may be a variant of any one of the
 30 xylosidases provided herein.

[0255] Arabinofuranosidase activity may be measured by chromogenic substrate 4-nitrophenyl alpha-L-arabinofuranoside (pNPA, Sigma-Aldrich N3641).

[0256] Any one of the compositions provided herein such as the enzyme blends/compositions of the disclosure can, for example, suitably comprise at least one polypeptide having arabinofuranosidase activity (e.g., L- α -arabinofuranosidase activity) such as L- α -arabinofuranosidases. The L- α -arabinofuranosidase may be, for example, Af43A, Fv43B, Pf51A, Pa51A, Fv51A, or a variant thereof.

[0257] The enzyme blends/compositions of the disclosure may optionally comprise one or more L- α -arabinofuranosidases in addition to or in place of the foregoing L- α -arabinofuranosidases. L- α -arabinofuranosidases (EC 3.2.1.55) from any suitable organism can be used as the additional L- α -arabinofuranosidases. Suitable L- α -arabinofuranosidases include, e.g., L- α -arabinofuranosidases of *A.oryzae* (Numan & Bhosle, *J. Ind. Microbiol. Biotechnol.* 2006, 33:247-260), *A.sojae* (Oshima *et al.* *J. Appl. Glycosci.* 2005, 52:261-265), *B.brevis* (Numan & Bhosle, *J. Ind. Microbiol. Biotechnol.* 2006, 33:247-260), *B.stearothermophilus* (Kim *et al.*, *J. Microbiol. Biotechnol.* 2004, 14:474-482), *B. breve* (Shin *et al.*, *Appl. Environ. Microbiol.* 2003, 69:7116-7123), *B.longum* (Margolles *et al.*, *Appl. Environ. Microbiol.* 2003, 69:5096-5103), *C.thermocellum* (Taylor *et al.*, *Biochem. J.* 2006, 395:31-37), *F.oxysporum* (Panagiotou *et al.*, *Can. J. Microbiol.* 2003, 49:639-644), *F.oxysporum f. sp. dianthi* (Numan & Bhosle, *J. Ind. Microbiol. Biotechnol.* 2006, 33:247-260), *G.stearothermophilus* T-6 (Shallom *et al.*, *J. Biol. Chem.* 2002, 277:43667-43673), *H.vulgare* (Lee *et al.*, *J. Biol. Chem.* 2003, 278:5377-5387), *P.chrysogenum* (Sakamoto *et al.*, *Biophys. Acta* 2003, 1621:204-210), *Penicillium sp.* (Rahman *et al.*, *Can. J. Microbiol.* 2003, 49:58-64), *P.cellulosa* (Numan & Bhosle, *J. Ind. Microbiol. Biotechnol.* 2006, 33:247-260), *R.pusillus* (Rahman *et al.*, *Carbohydr. Res.* 2003, 338:1469-1476), *S.chartreusis*, *S.thermoviolacus*, *T.ethanolicus*, *T.xylanilyticus* (Numan & Bhosle, *J. Ind. Microbiol. Biotechnol.* 2006, 33:247-260), *T.fusca* (Tuncer and Ball, *Folia Microbiol.* 2003, (Praha) 48:168-172), *T.maritima* (Miyazaki, *Extremophiles* 2005, 9:399-406), *Trichoderma sp.* SY (Jung *et al.* *Agric. Chem. Biotechnol.* 2005, 48:7-10), *A.kawachii* (Koseki *et al.*, *Biochim. Biophys. Acta* 2006, 1760:1458-1464), *F.oxysporum f. sp. dianthi* (Chacon-Martinez *et al.*, *Physiol.Mol. Plant Pathol.* 2004,64:201-208), *T.xylanilyticus* (Debeche *et al.*, *Protein Eng.* 2002, 15:21-28), *H.insolens*, *M.giganteus* (Sorensen *et al.*, *Biotechnol. Prog.* 2007, 23:100-107), or *R.sativus* (Kotake *et al.* *J. Exp. Bot.* 2006, 57:2353-2362). A polypeptide having arabinofuranosidase activity may be a variant of any one of the arabinofuranosidases described herein.

[0258] In some aspects of any one of the compositions described herein, the at least one polypeptide having endoglucanase activity comprises *T. reesei* EG1 (or a variant thereof) and/or *T. reesei* EG2 (or a variant thereof). In some aspects of any one of the compositions described herein, the at least one polypeptide having cellobiohydrolase (“CBH”) activity comprises *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof. In some aspects of any one of the compositions described herein, the at least one polypeptide having β -glucosidase activity comprises Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, and/or Tn3B. In some aspects of any one of the compositions described herein, the at least one polypeptide having β -glucosidase activity comprises Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B, and/or a variant thereof. In some aspects of any one of the compositions described herein, the at least one polypeptide having xylanase activity comprises *T. reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, and/or AfuXyn5. In some aspects of any one of the compositions described herein, the at least one polypeptide having xylanase activity comprises *T. reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, AfuXyn5, and/or a variant thereof. In some aspects of any one of the compositions described herein, the at least one polypeptide having β -xylosidase activity is a Group 1 β -xylosidase or a Group 2 β -xylosidase, wherein the Group 1 β -xylosidase comprises Fv3A, Fv43A, or a variant thereof, and the Group 2 β -xylosidase comprises Pf43A, Fv43D, Fv39A, Fv43E, Fo43A, Fv43B, Pa51A, Gz43A, *T. reesei* Bxl1, or a variant thereof. In some aspects, the at least one polypeptide having β -xylosidase activity comprises *F. verticillioides* Fv3A, *F. verticillioides* Fv43D, or a variant thereof. In some aspects of any one of the compositions described herein, the at least one polypeptide having L- α -arabinofuranosidase activity comprises Af43A, Fv43B, Pf51A, Pa51A, and/or Fv51A. In some aspects of any one of the compositions described herein, the at least one polypeptide having L- α -arabinofuranosidase activity comprises Af43A, Fv43B, Pf51A, Pa51A, Fv51A, and/or a variant thereof.

Whole cellulase

[0259] Any of the compositions provided here such as enzyme blends/compositions of the disclosure may comprise whole cellulase.

[0260] As used herein, a “whole cellulase” refers to either a naturally occurring or a non-naturally occurring cellulase-containing composition comprising at least 3 different enzyme

types: (1) an endoglucanase, (2) a cellobiohydrolase, and (3) a β -glucosidase, or comprising at least 3 different enzymatic activities: (1) an endoglucanase activity, which catalyzes the cleavage of internal β -1,4 linkages, resulting in shorter glucooligosaccharides, (2) a cellobiohydrolase activity, which catalyzes an “exo”-type release of cellobiose units (β -1,4 glucose-glucose disaccharide), and (3) a β -glucosidase activity, which catalyzes the release of glucose monomer from short cellooligosaccharides (*e.g.*, cellobiose). The whole cellulase may comprise at least one polypeptide having endoglucanase activity (*e.g.*, EG2 (or a variant thereof) and/or EG4 (or a variant thereof)), at least one polypeptide having cellobiohydrolase activity (*e.g.*, CBH1 (or a variant thereof) and/or CBH2 (or a variant thereof)), and at least one polypeptide having β -glucosidase activity (*e.g.*, Bgl1 or a variant thereof).

[0261] A “naturally occurring cellulase-containing” composition is one produced by a naturally occurring source, which comprises one or more cellobiohydrolase-type, one or more endoglucanase-type, and one or more β -glucosidase-type components or activities, wherein each of these components or activities is found at the ratio and level produced in nature, untouched by the human hand. Accordingly, a naturally occurring cellulase-containing composition is, for example, one that is produced by an organism unmodified with respect to the cellulolytic enzymes such that the ratio or levels of the component enzymes are unaltered from that produced by the native organism in nature. A “non-naturally occurring cellulase-containing composition” refers to a composition produced by: (1) combining component cellulolytic enzymes either in a naturally occurring ratio or a non-naturally occurring, *i.e.*, altered, ratio; or (2) modifying an organism to overexpress or underexpress one or more cellulolytic enzymes; or (3) modifying an organism such that at least one cellulolytic enzyme is deleted. A “non-naturally occurring cellulase containing” composition can also refer to a composition resulting from adjusting the culture conditions for a naturally-occurring organism, such that the naturally-occurring organism grows under a non-native condition, and produces an altered level or ratio of enzymes. Accordingly, in some embodiments, the whole cellulase preparation of the present disclosure can have one or more EGs and/or CBHs and/or β -glucosidases deleted and/or overexpressed.

[0262] In some aspects, there is provided a non-naturally occurring composition comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, endoglucanase IV polypeptide such as *T. reesei* Eg4 polypeptide or a variant thereof) or a non-naturally occurring composition comprising a polypeptide having GH61/ endoglucanase activity (*e.g.*,

whole cellulase enriched with endoglucanase IV polypeptide such as *T. reesei* Eg4 polypeptide or a variant thereof), wherein the composition further comprises a whole cellulase, at least 1 polypeptide having endoglucanase activity (*e.g.*, at least 2, 3, 4, or 5 polypeptides having endoglucanase activity), at least 1 polypeptide having cellobiohydrolase activity (*e.g.*, at least 2, 3, 4, or 5 polypeptides having cellobiohydrolase activity), at least 1 polypeptide having glucosidase activity (*e.g.*, β -glucosidase) (*e.g.*, at least 2, 3, 4, or 5 polypeptides having β -glucosidase activity), at least 1 polypeptide having xylanase activity (*e.g.*, at least 2, 3, 4, or 5 polypeptides having xylanase activity), at least 1 polypeptide having xylosidase activity (*e.g.*, β -xylosidase) (*e.g.*, at least 2, 3, 4, or 5 polypeptides having β -xylosidase activity), and/or at least 1 polypeptide having arabinofuranosidase activity (*e.g.*, L- α -arabinofuranosidase) (*e.g.*, at least 2, 3, 4, or 5 polypeptides having L- α -arabinofuranosidase activity). The polypeptides having various enzyme activities are described above.

[0263] In some aspects, the whole cellulase comprises at least 1 polypeptide having endoglucanase activity such as *T. reesei* EG1, *T. reesei* EG2, or a variant thereof. In some aspects, the whole cellulase comprises at least one polypeptide having cellobiohydrolase activity such as *T. reesei* CBH1, *T. reesei* CBH2, or a variant thereof. In some aspects, the whole cellulase comprises at least 1 polypeptide having β -glucosidase activity such as Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A, Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B, or a variant thereof.

[0264] In the present disclosure, a whole cellulase preparation can be from any microorganism that is capable of hydrolyzing a cellulosic material. In some embodiments, the whole cellulase preparation is a fungal or bacterial whole cellulase. For example, the whole cellulase preparation can be from an *Acremonium*, *Aspergillus*, *Chrysosporium*, *Emericella*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Scytalidium*, *Thielavia*, *Tolypocladium*, *Trichoderma*, or yeast species.

[0265] The whole cellulase preparation may be, *e.g.*, an *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, or *Aspergillus oryzae* whole cellulase. Moreover, the whole cellulase preparation may be a *Fusarium bac-tridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium*

sambucinum, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, or *Fusarium venenatum* whole cellulase preparation. The whole cellulase preparation may also be a *Chrysosporium lucknowence*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*,

5 *Neurospora crassa*, *Penicillium purpurogenum*, *Penicillium funiculosum*, *Scytalidium thermophilum*, or *Thielavia terrestris* whole cellulase preparation. The whole cellulase preparation may also be a *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* (e.g., RL-P37 (Sheir-Neiss G et al. Appl. Microbiol. Biotechnology, 1984, 20, pp.46-53), QM9414 (ATCC No. 26921), NRRL 15709, ATCC
10 13631, 56764, 56466, 56767), or a *Trichoderma viride* (e.g., ATCC 32098 and 32086) whole cellulase preparation.

[0266] The whole cellulase preparation can be integrated strain *T.reesei* H3A or H3A/Eg4 #27 (as described in the Examples herein) preparation.

[0267] The whole cellulase preparation can suitably be a *T.reesei* RutC30 whole cellulase
15 preparation, which is available from the American Type Culture Collection as *T.reesei* ATCC 56765. For example, the whole cellulase preparation can also suitably be a whole cellulase of *P. funiculosum*, which is available from the American Type Culture Collection as *P. funiculosum* ATCC Number: 10446.

[0268] The whole cellulase preparation can also be obtained from commercial sources.
20 Examples of commercial cellulase preparations suitable for use in the methods and compositions of the present disclosure include, for example, CELLUCLAST™ and Cellic™ (Novozymes A/S) and LAMINEX™ BG, IndiAge™ 44L, Primafast™ 100, Primafast™ 200, Spezyme™ CP, Accellerase® 1000 and Accellerase® 1500 (Danisco US. Inc., Genencor).

[0269] Suitable whole cellulase preparations can be made using any known microorganism
25 cultivation methods, especially fermentation, resulting in the expression of enzymes capable of hydrolyzing a cellulosic material. As used herein, “fermentation” refers to shake flask cultivation, small- or large-scale fermentation, such as continuous, batch, fed-batch, or solid state fermentations in laboratory or industrial fermenters performed in a suitable medium and under conditions that allow the cellulase and/or enzymes of interest to be expressed and/or
30 isolated. Generally the microorganism is cultivated in a cell culture medium suitable for production of enzymes capable of hydrolyzing a cellulosic material. The cultivation takes place in a nutrient medium comprising carbon and nitrogen sources and inorganic salts, using

known procedures and variations. Culture media, temperature ranges and other conditions for growth and cellulase production are known. As a non-limiting example, a typical temperature range for the production of cellulases by *T. reesei* is 24°C to 28°C.

[0270] The whole cellulase preparation can be used as it is produced by fermentation with no or minimal recovery and/or purification. In that sense, the whole cellulase preparation can be used in a whole broth formulation. For example, once cellulases are secreted into the cell culture medium, the cell culture medium containing the cellulases can be used directly. The whole cellulase preparation can comprise the unfractionated contents of fermentation material, including the spent cell culture medium, extracellular enzymes and cells. On the other hand, the whole cellulase preparation can also be subject to further processing in a number of routine steps, *e.g.*, precipitation, centrifugation, affinity chromatography, filtration, or the like. For example, the whole cellulase preparation can be concentrated, and then used without further purification. The whole cellulase preparation can, *e.g.*, be formulated to comprise certain chemical agents that decrease cell viability or kill the cells after fermentation. The cells can for example be lysed or permeabilized using known methods.

[0271] The endoglucanase activity of the whole cellulase preparation can be determined using carboxymethyl cellulose (CMC) as a substrate. A suitable assay measures the production of reducing ends created by the enzyme mixture acting on CMC wherein 1 unit is the amount of enzyme that liberates 1 µmol of product/min (Ghose, T. K., *Pure & Appl. Chem.* 1987, 59, pp. 257-268).

[0272] The whole cellulase may be enriched with a polypeptide having GH61/endoglucanase activity, *e.g.*, an EG IV-enriched (such as, *e.g.*, enriched with *T. reesei* Eg4 polypeptide or a variant thereof) cellulase. The EG IV-enriched whole cellulase generally comprises an EG IV polypeptide (such as, *e.g.*, *T. reesei* Eg4 polypeptide or a variant thereof) and a whole cellulase preparation. The EG IV-enriched whole cellulase compositions can be produced by recombinant means. For example, such a whole cellulase preparation can be achieved by expressing an EG IV in a microorganism capable of producing a whole cellulase. The EG IV-enriched whole cellulase composition can also, *e.g.*, comprise a whole cellulase preparation and an EG IV (such as, *e.g.*, *T. reesei* Eg4 polypeptide or a variant thereof). For instance, the EG IV-enriched (*e.g.*, enriched with *T. reesei* Eg4 polypeptide or a variant thereof) whole cellulase composition can suitably comprise at least 0.1 wt.%, 1 wt.%, 2 wt.%, 5 wt.%, 7 wt.%, 10 wt.%, 15 wt.% or 20 wt.%, and up to 25 wt.%, 30 wt.%, 35 wt.%, 40

wt.%, or 50 wt.% EG IV based on the total weight of proteins in that blend/composition.

[0273] The whole cellulase can be a β -glucosidase-enriched cellulase. The β -glucosidase-enriched whole cellulase generally comprises a β -glucosidase and a whole cellulase preparation. The β -glucosidase-enriched whole cellulase compositions can be produced by recombinant means. For example, such a whole cellulase preparation can be achieved by expressing a β -glucosidase in a microorganism capable of producing a whole cellulase. The β -glucosidase-enriched whole cellulase composition can also, e.g., comprise a whole cellulase preparation and a β -glucosidase. For instance, the β -glucosidase-enriched whole cellulase composition can suitably comprise at least 0.1 wt.%, 1 wt.%, 2 wt.%, 5 wt.%, 7 wt.%, 10 wt.%, 15 wt.% or 20 wt.%, and up to 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, or 50 wt.% β -glucosidase based on the total weight of proteins in that blend/composition.

[0274] Certain fungi produce complete cellulase systems, including exo-cellobiohydrolases or CBH-type cellulases, endoglucanases or EG-type cellulases and β -glucosidase or BG-type cellulases (Schulein, 1988). However, sometimes these systems lack CBH-type cellulases, e.g., bacterial cellulases also typically include little or no CBH-type cellulases. In addition, it has been shown that the EG components and CBH components synergistically interact to more efficiently degrade cellulose. See, e.g., Wood, 1985. The different components, i.e., the various endoglucanases and exocellobiohydrolases in a multi-component or complete cellulase system, generally have different properties, such as isoelectric point, molecular weight, degree of glycosylation, substrate specificity and enzymatic action patterns.

[0275] In some aspects, the cellulase is used as is produced by fermentation with no or minimal recovery and/or purification. For example, once cellulases are secreted by a cell into the cell culture medium, the cell culture medium containing the cellulases can be used. In some aspects, the whole cellulase preparation comprises the unfractionated contents of fermentation material, including cell culture medium, extracellular enzymes and cells. Alternatively, the whole cellulase preparation can be processed by any convenient method, e.g., by precipitation, centrifugation, affinity, filtration or any other method known in the art. In some aspects, the whole cellulase preparation can be concentrated, for example, and then used without further purification. In some aspects, the whole cellulase preparation comprises chemical agents that decrease cell viability or kills the cells. In some aspects, the cells are lysed or permeabilized using methods known in the art.

[0276] A composition is provided comprising a polypeptide having GH61/endoglucanase

activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and further comprising at least one cellulase polypeptide and/or at least one hemicellulase polypeptide, wherein the cellulase polypeptide and/or the hemicellulase polypeptide is heterologous to the host cell expressing the cellulase polypeptide and/or the hemicellulase polypeptide. In some aspects, there is provided a

- 5 composition comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and further comprising at least 1 cellulase polypeptide and/or at least 1 hemicellulase polypeptide, wherein the cellulase polypeptide and/or the hemicellulase polypeptide is expressed from a host cell, and wherein cellulase polypeptide and/or a hemicellulase polypeptide is endogenous to the host cell. The cellulase polypeptide may
- 10 comprise a polypeptide having endoglucanase activity (*e.g.*, *T. reesei* EG1 or a variant thereof, *T. reesei* EG2 or a variant thereof), a polypeptide having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1, *A. fumigatus* 7A, 7B, *C. globosum* 7A, 7B, *T. terrestris* 7A, 7B, *T. reesei* CBH2, *T. terrestris* 6A, *S. thermophile* 6A, 6B, or a variant thereof), or a polypeptide having β -glucosidase activity (*e.g.*, Fv3C, Pa3D, Fv3G, Fv3D, Tr3A, Tr3B, Te3A, An3A,
- 15 Fo3A, Gz3A, Nh3A, Vd3A, Pa3G, Tn3B, or a variant thereof). The hemicellulase polypeptide may comprise a polypeptide having xylanase activity (*e.g.*, *T. reesei* Xyn3, *T. reesei* Xyn2, AfuXyn2, AfuXyn5, or a variant thereof), a having β -xylosidase activity (*e.g.*, Fv3A, Fv43A, Pf43A, Fv43D, Fv39A, Fv43E, Fo43A, Fv43B, Pa51A, Gz43A, *T. reesei* Bxl1, or a variant thereof), or a polypeptide having L- α -arabinofuranosidase activity (*e.g.*,
- 20 Af43A, Fv43B, Pf51A, Pa51A, Fv51A, or a variant thereof).

[0277] In some aspects, the composition is from fermentation broth. The composition may be from the fermentation broth of a strain, wherein a nucleic acid encoding a polypeptide having GH61/ endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is heterologous to the host cell expressing the polypeptide having GH61/endoglucanase activity (*e.g.*,

- 25 integrated into the strain or expressed from a vector in the host strain). The composition may be from the fermentation broth of an integrated strain (*e.g.*, H3A/Eg4, #27 as in Examples).

[0278] The composition comprising a polypeptide having GH61/ endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) may comprise whole cellulase. Thus, a composition is provided (*e.g.*, a non-naturally occurring composition) comprising *T. reesei* Eg4 (or a

30 variant thereof), *T. reesei* Bgl1 (or a variant thereof), *T. reesei* xyn3 (or a variant thereof), Fv3A (or a variant thereof), Fv43D (or a variant thereof), and Fv51A (or a variant thereof).

[0279] In some aspects, the composition comprises isolated *T. reesei* Eg4. In some aspects,

the composition comprises at least one (at least 2, 3, 4, or 5) of isolated *T. reesei* Bgl1, isolated *T. reesei* xyn3, isolated Fv3A, isolated Fv43D, and isolated Fv51A.

[0280] In some aspects, the composition is from fermentation broth. In some aspects, the composition is from the fermentation broth of an integrated strain (*e.g.*, H3A/Eg4, #27 as described herein in the Examples). The *T. reesei* Eg4 or the nucleic acid encoding *T. reesei* Eg4 may be heterologous to the host cell expressing *T. reesei* Eg4. At least one nucleic acid encoding *T. reesei* Bgl1, *T. reesei* xyn3, Fv3A, Fv43D, Fv51A, or a variant thereof may be heterologous to the host cell such as the host cell expressing *T. reesei* Eg4. In some aspects, at least one nucleic acid encoding *T. reesei* Bgl1, *T. reesei* xyn3, Fv3A, Fv43D, Fv51A, or a variant thereof is endogenous to the host cell such as the host cell expressing *T. reesei* Eg4.

[0281] Regarding any of the compositions described above, varying amounts of the polypeptide(s) included in the compositions are described below in “Amount of component(s) in compositions” section.

Amount of component(s) in compositions

[0282] A non-naturally occurring composition comprising a polypeptide having GH61/endoglucanase activity (or a non-naturally occurring composition comprising whole cellulase comprising a polypeptide having GH61/endoglucanase activity) provided herein may comprise various components as described herein, wherein each component is present in the composition in various amount.

[0283] In some aspects of any one of the compositions or methods provided herein, the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is present in the composition in an amount sufficient to increase the yield of fermentable sugar(s) from hydrolysis of biomass material (*e.g.*, by at least about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%) compared to the yield in the absence of the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof). Any one of the compositions or methods provided herein, the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) may be present in the composition in an amount sufficient to reduce the viscosity of a biomass mixture during hydrolysis of a biomass material (*e.g.*, by at least about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%) compared to the viscosity of the biomass mixture during hydrolysis in the absence of the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof). The composition may further

comprise at least 1 polypeptide having endoglucanase activity, at least 1 polypeptide having cellobiohydrolase activity, at least 1 polypeptide having β -glucosidase activity, at least 1 polypeptide having xylanase activity, at least 1 polypeptide having β -xylosidase activity, at least 1 polypeptide having L- α -arabinofuranosidase activity, and/or whole cellulase, or a mixture thereof. The amount of polypeptide(s) having endoglucanase activity, the amount of polypeptide(s) having cellobiohydrolase activity, the amount of polypeptide(s) having β -glucosidase activity, the amount of polypeptide(s) having xylanase activity, the amount of polypeptide(s) having β -xylosidase activity, the amount of polypeptide(s) having L- α -arabinofuranosidase activity, or the amount of whole cellulase is sufficient to increase the yield of fermentable sugar(s) from hydrolysis of biomass material (*e.g.*, by at least about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%) compared to the yield in the absence of the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof), the polypeptide(s) having endoglucanase activity, the polypeptide(s) having cellobiohydrolase activity, the polypeptide(s) having β -glucosidase activity, the polypeptide(s) having xylanase activity, the polypeptide(s) having β -xylosidase activity, the polypeptide(s) having L- α -arabinofuranosidase activity, or the whole cellulase. In some aspects, the amount of polypeptide(s) having endoglucanase activity, the amount of polypeptide(s) having cellobiohydrolase activity, the amount of polypeptide(s) having β -glucosidase activity, the amount of polypeptide(s) having xylanase activity, the amount of polypeptide(s) having β -xylosidase activity, the amount of polypeptide(s) having L- α -arabinofuranosidase activity, or the amount of whole cellulase is sufficient to reduce the viscosity of a biomass mixture during hydrolysis of a biomass material (*e.g.*, by at least about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%) compared to the viscosity of a biomass mixture in the absence of the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof), the polypeptide(s) having endoglucanase activity, the polypeptide(s) having cellobiohydrolase activity, the polypeptide(s) having β -glucosidase activity, the polypeptide(s) having xylanase activity, the polypeptide(s) having β -xylosidase activity, the polypeptide(s) having L- α -arabinofuranosidase activity, or the whole cellulase.

[0284] A polypeptide having GH61/endoglucanase activity (such as EG IV including *T. reesei* Eg4 polypeptide or a variant thereof) may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an

amount that is at least about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of proteins in the composition. In some aspects, a polypeptide having GH61/endoglucanase activity (such as EG IV including, *e.g.*, *T. reesei* Eg4 polypeptide or a variant thereof) may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that is no more than about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, or 80 wt.% of the total weight of proteins in the composition. A polypeptide having GH61/endoglucanase activity (such as EG IV including, *e.g.*, *T. reesei* Eg4 polypeptide or a variant thereof) may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that has a range having upper limit and lower limit. For example, lower limit for a polypeptide having GH61/endoglucanase activity is about any of 0.01 wt.%, 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of proteins in the composition. Upper limit for a polypeptide having GH61/endoglucanase activity may be about any of 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.% or 70 wt.% of the total weight of proteins in the composition. In some aspects, a polypeptide having GH61/endoglucanase activity (such as EG IV including, *e.g.*, *T. reesei* Eg4 polypeptide or a variant thereof) may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that is about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, or 80 wt.% of the total weight of proteins in the composition. The polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) may be present in about 10 wt.% or 12 wt.% of the total weight of proteins in the composition. The composition may have at least two polypeptides having endoglucanase activity (*e.g.*, *T. reesei* Eg4, *T. reesei* Eg1, and/or *T. reesei* Eg2, or a variant thereof), where the total amount of polypeptides having endoglucanase activity is about 0.1 to about 50 wt.% (*e.g.*, about 0.5 to about 45 wt.%, about 1 to about 30 wt.%, about 2 to about 20 wt.%, about 5 to about 20

wt.%, or about 8 to about 15 wt.%) of the total weight of proteins in the composition. The polypeptide having GH61/endoglucanase activity may be heterologous or endogenous to the host cell expressing the polypeptide having GH61/endoglucanase activity. The polypeptide having GH61/ endoglucanase activity included in the composition may be isolated.

5 [0285] In some aspects, the enzyme composition (*e.g.*, the enzyme composition) described herein is whole cellulase composition comprising a polypeptide having GH61/endoglucanase activity. In some aspects, a polypeptide having GH61/endoglucanase activity (such as EG IV including, *e.g.*, *T. reesei* Eg4 polypeptide or a variant thereof) may be present in an amount that is at least about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%,
10 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of the whole cellulase. In some aspects, a polypeptide having GH61/endoglucanase activity (such as EG IV including, *e.g.*, *T. reesei* Eg4 polypeptide or a variant thereof) may be present in an amount that is no more than about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%,
15 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, or 80 wt.% of the total weight of the whole cellulase. In some aspects, a polypeptide having GH61/endoglucanase activity (such as EG IV including, *e.g.*, *T. reesei* Eg4 polypeptide or a variant thereof) may be present in an amount that has a lower limit of about any of 0.01 wt.%, 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of the whole cellulase and a upper limit of about any of 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.% or 70 wt.% of the total weight of the whole cellulase. In some aspects, a polypeptide having GH61/endoglucanase activity (such as EG IV including, *e.g.*, *T. reesei* Eg4
25 polypeptide or a variant thereof) may be present in an amount that is about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 13 wt.%, 14 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, or 80 wt.% of the total weight of the whole cellulase. In some aspects, a polypeptide having GH61/endoglucanase activity (such as EG
30 IV including, *e.g.*, *T. reesei* Eg4 polypeptide or a variant thereof) is present in an amount that is about 10 wt.% or 12 wt.% of the total weight of the whole cellulase.

[0286] In some aspects, any of the compositions provided herein may comprise at least one

polypeptide having endoglucanase activity (e.g., in addition to a polypeptide having GH61/endoglucanase activity) including *T. reesei* Eg1 or a variant thereof and/or *T. reesei* Eg2 or a variant thereof. In some aspects, the total amount of the polypeptide(s) having endoglucanase activity may be present in any of the compositions described herein (such as

5 in any of the enzyme blends/compositions provided herein) in an amount that is at least about 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of proteins in the composition. In some aspects, the total amount of the polypeptide(s) having endoglucanase activity may be present in any of the compositions described herein

10 (such as in any of the enzyme blends/compositions provided herein) in an amount that is no more than about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, or 80 wt.% of the total weight of proteins in the composition. In some aspects, the total amount of the polypeptide(s) having

15 endoglucanase activity may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that has a range having upper limit and lower limit. For example, lower limit for the total amount of the polypeptide(s) having endoglucanase activity is about any of 0.01 wt.%, 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 12 wt.%, 15 wt.%, 20 wt.%,

20 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of proteins in the composition. Upper limit for the total amount of the polypeptide(s) having endoglucanase activity may be about any of 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.% or 70 wt.% of the total weight of proteins in the composition. In some aspects, the total amount of the polypeptide(s) having endoglucanase

25 activity may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that is about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, or 80 wt.% of the total weight of proteins in the composition.

30 **[0287]** In some aspects, any of the compositions provided herein may comprise one or more polypeptide with various enzyme activity, such as polypeptide(s) having cellobiohydrolase activity, polypeptide(s) having glucosidase activity (e.g., β -glucosidase), polypeptide(s)

having xylanase activity, polypeptide(s) having xylosidase activity, and/or polypeptide(s) having arabinofuranosidase activity. In some aspects, there may be multiple polypeptides having the same enzyme activity. Each of the polypeptides mentioned above (or the total amount of the polypeptides having a specific enzyme activity, *e.g.*, total amount of the

5 polypeptides having cellobiohydrolase activity, glucosidase activity (*e.g.*, β -glucosidase), xylanase activity, xylosidase activity, or arabinofuranosidase activity) may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that is at least about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25

10 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of proteins in the composition. In some aspects, each of the polypeptides mentioned above (or the total amount of the polypeptides having a specific enzyme activity, *e.g.*, total amount of the polypeptides having cellobiohydrolase activity, glucosidase activity (*e.g.*, β -glucosidase), xylanase activity, xylosidase activity, or arabinofuranosidase activity) may be no more than about any of 1

15 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, or 80 wt.% of the total weight of proteins in the composition. Each of the polypeptides mentioned above (or the total amount of the polypeptides having a specific enzyme activity, *e.g.*, total amount of the polypeptides having

20 cellobiohydrolase activity, glucosidase activity (*e.g.*, β -glucosidase), xylanase activity, xylosidase activity, or arabinofuranosidase activity) may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that has a range having upper and lower limits. For example, lower limit for the total amount of the polypeptide(s) having endoglucanase activity is about any of 0.01

25 wt.%, 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of proteins in the composition. Upper limit may be about any of 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.% or 70 wt.% of the total weight of proteins in the composition. In some aspects, each of the polypeptides

30 mentioned above (or the total amount of the polypeptides having a specific enzyme activity, *e.g.*, total amount of the polypeptides having cellobiohydrolase activity, glucosidase activity (*e.g.*, β -glucosidase), xylanase activity, xylosidase activity, or arabinofuranosidase activity)

may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that is about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, or 80 wt.% of the total weight of proteins in the composition.

[0288] In some aspects, any of the compositions provided herein may further comprise whole cellulase. The whole cellulase may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that is at least about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, 80 wt.%, 85 wt.%, 90 wt.%, or 95 wt.% of the total weight of proteins in the composition. The whole cellulase may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that is no more than about any of 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, 80 wt.%, 85 wt.%, 90 wt.%, or 95 wt.% of the total weight of proteins in the composition. The whole cellulase may be present in any of the compositions described herein (such as in any of the enzyme blends/compositions provided herein) in an amount that is about any of 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, 80 wt.%, 85 wt.%, 90 wt.%, or 95 wt.% of the total weight of proteins in the composition.

[0289] In some aspects of any one of the compositions or methods provided herein, the polypeptide having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1, *T. reesei* CBH2, or a variant thereof) is present in an amount that is about 0.1 to about 70 wt.% (*e.g.*, about 0.5 to about 60 wt.%, about 5 to about 70 wt.%, about 10 to about 60 wt.%, about 20 to about 50 wt.%, or about 30 to about 50 wt.%) of the total weight of proteins in the composition. In some aspects, the composition has at least two polypeptides having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1 (or a variant thereof) and *T. reesei* CBH2 (or a variant thereof)), wherein the total amount of polypeptides having cellobiohydrolase activity is about 0.1 to about 70 wt.% (*e.g.*, about 0.5 to about 60 wt.%, about 5 to about 70 wt.%, about 10 to about 60 wt.%, about 20 to about 50 wt.%, or about 30 to about 50 wt.%) of the total weight of

proteins in the composition. The polypeptide having cellobiohydrolase activity may be expressed from a nucleic acid heterologous or endogenous to the host cell. In some aspects, the polypeptide having cellobiohydrolase activity included in the composition is isolated.

[0290] In some aspects of any one of the compositions or methods provided herein, the

5 polypeptide having β -glucosidase activity (*e.g.*, an Fv3C, a Pa3D, an Fv3G, an Fv3D, a Tr3A, a Tr3B, a Te3A, an An3A, an Fo3A, a Gz3A, an Nh3A, a Vd3A, a Pa3G, a Tn3B, or a variant thereof) is present in an amount that is about 0.1 to about 50 wt.% (*e.g.*, about 0.5 to about 40 wt.%, about 1 to about 30 wt.%, about 2 to about 20 wt.%, about 5 to about 20 wt.%, or about 8 to about 15 wt.%) of the total weight of proteins in the composition. In some
10 aspects, the composition has at least two polypeptides having β -glucosidase activity, wherein the total amount of polypeptides having β -glucosidase activity is about 0.1 to about 50 wt.% (*e.g.*, about 0.5 to about 40 wt.% about 1 to about 30 wt.%, about 2 to about 20 wt.%, about 5 to about 20 wt.%, or about 8 to about 15 wt.%) of the total weight of proteins in the composition. The polypeptide having β -glucosidase activity may be expressed from a nucleic
15 acid heterologous or endogenous to the host cell. In some aspects, the polypeptide having β -glucosidase activity included in the composition is isolated.

[0291] Any one of the compositions or methods provided herein, the polypeptide having xylanase activity (*e.g.*, *T. reesei* Xyn3, *T. reesei* Xyn2, an AfuXyn2, an AfuXyn5, or a variant thereof) may be present in an amount that is about 0.1 to about 50 wt.% (*e.g.*, about 0.5 to
20 about 40 wt.%, about 1 to about 40 wt.%, about 4 to about 30 wt.%, about 5 to about 20 wt.%, or about 8 to about 15 wt.%) of the total weight of proteins in the composition. The composition may have at least 2 polypeptides having xylanase activity, wherein the total amount of polypeptides having xylanase activity is about 0.1 to about 50 wt.% (*e.g.*, about 0.5 to about 40 wt.%, about 1 to about 40 wt.%, about 4 to about 30 wt.%, about 5 to about
25 20 wt.%, or about 8 to about 15 wt.%) of the total weight of proteins in the composition. The polypeptide having xylanase activity may be expressed from a nucleic acid heterologous or endogenous to the host cell. The polypeptide having xylanase activity included in the composition may be isolated.

[0292] Any one of the compositions or methods provided herein, the polypeptide having L-
30 α -arabinofuranosidase activity (*e.g.*, an Af43A, an Fv43B, a Pf51A, a Pa51A, an Fv51A, or a variant thereof) may be present in an amount that is about 0.1 to about 50 wt.% (*e.g.*, about 0.5 to about 45 wt.%, about 1 to about 40 wt.%, about 2 to about 30 wt.%, about 4 to about

20 wt.%, or about 5 to about 15 wt.%) of the total weight of enzymes in the composition. The composition may have at least 2 polypeptides having L- α -arabinofuranosidase activity, wherein the total amount of polypeptides having L- α -arabinofuranosidase activity is about 0.1 to about 50 wt.% (e.g., about 0.5 to about 45 wt.%, about 1 to about 40 wt.%, about 2 to about 30 wt.%, about 4 to about 20 wt.%, or about 5 to about 15 wt.%) of the total weight of proteins in the composition. The polypeptide having L- α -arabinofuranosidase activity may be expressed from a nucleic acid heterologous or heterologous to the host cell. The polypeptide having L- α -arabinofuranosidase activity included in the composition may be isolated.

[0293] Any one of the compositions or methods provided herein, the polypeptide having β -xylosidase activity (e.g., Fv3A, Fv43A, a Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fo43A, an Fv43B, a Pa51A, a Gz43A, a *T. reesei* Bx11, or a variant thereof) may be present in an amount that is about 0.1 to about 50 wt.% (e.g., about 0.5 to about 45 wt.%, about 1 to about 40 wt.%, about 4 to about 35 wt.%, about 5 to about 25 wt.%, or about 5 to about 20 wt.%) of the total weight of enzymes in the composition. The composition may have at least 2 polypeptides having β -xylosidase activity, wherein the total amount of polypeptides having β -xylosidase activity is about 0.1 to about 50 wt.% (e.g., about 0.5 to about 45 wt.%, about 1 to about 40 wt.%, about 4 to about 35 wt.%, about 5 to about 25 wt.%, or about 5 to about 20 wt.%) of the total weight of proteins in the composition. The polypeptide having β -xylosidase activity may be expressed from a nucleic acid heterologous or endogenous to the host cell. The polypeptide having β -xylosidase activity included in the composition may be isolated.

[0294] Any one of the compositions or methods provided herein, the whole cellulase in the composition may be about 0.1 to about 100 wt.% (e.g., about 1 to about 95 wt.%, about 5 to about 90 wt.%, about 10 to about 85 wt.%, about 20 to about 80 wt.%, or about 30 to about 75 wt.%) of the total weight of proteins in the composition. The whole cellulase may comprise at least 1 polypeptide having endoglucanase activity (such as *T. reesei* Eg4 or a variant thereof, *T. reesei* Eg1 or a variant thereof, *T. reesei* Eg2 or a variant thereof) expressed from a nucleic acid heterologous or endogenous to the host cell. The whole cellulase may comprise at least 1 polypeptide having cellobiohydrolase activity (e.g., *T. reesei* CBH1 or a variant thereof, *T. reesei* CBH2 or a variant thereof) expressed from a nucleic acid heterologous or endogenous to the host cell. The whole cellulase may comprise at least one polypeptide having β -glucosidase activity (e.g., an Fv3C, a Pa3D, an Fv3G, an Fv3D, a Tr3A, a Tr3B, a

Te3A, an An3A, an Fo3A, a Gz3A, an Nh3A, a Vd3A, a Pa3G, a Tn3B, or a variant thereof) expressed from a nucleic acid heterologous or endogenous to the host cell.

[0295] In some aspects, the composition of the invention is capable of converting a biomass material into fermentable sugar(s) (*e.g.*, glucose, xylose, arabinose, and/or cellobiose). In some aspects, the composition is capable of achieving at least 0.1 (*e.g.*, 0.1 to 0.4) fraction product as determined by the calcofluor assay.

[0296] In some aspects, the composition comprises the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and further comprises at least one cellulase polypeptide and/or at least one hemicellulase polypeptide, wherein the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one cellulase polypeptide and/or at least one hemicellulase polypeptide are mixed together before contacting a biomass material.

[0297] In some aspects, the composition comprises a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and further comprises at least one cellulase polypeptide and/or at least one hemicellulase polypeptide, wherein the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one cellulase polypeptide and/or at least one hemicellulase polypeptide are added to a biomass material at different times (*e.g.*, a polypeptide having GH61/endoglucanase activity is added to a biomass material before or after the at least one cellulase polypeptide and/or at least one hemicellulase polypeptide is added to the biomass material).

[0298] In some aspects, the composition comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is a mixture comprising a biomass material, *e.g.*, the composition is a hydrolysis mixture, a fermentation mixture, or a saccharification mixture. Such mixture may further include fermentable sugar(s).

Other components

[0299] The enzyme compositions of the disclosure may suitably further comprise 1 or more accessory proteins. Examples of accessory proteins include, without limitation, mannanases (*e.g.*, endomannanases, exomannanases, and β -mannosidases), galactanases (*e.g.*, endo- and exo-galactanases), arabinases (*e.g.*, endo-arabinases and exo-arabinases), ligninases, amylases, glucuronidases, proteases, esterases (*e.g.*, ferulic acid esterases, acetyl xylan esterases, coumaric acid esterases or pectin methyl esterases), lipases, other glycoside hydrolases, xyloglucanases, CIP1, CIP2, swollenins, expansins, and cellulose disrupting

proteins. For example, the cellulose disrupting proteins are cellulose binding modules.

Methods and processes

[0300] The disclosure provides methods and processes for biomass saccharification, using enzymes, enzyme blends/compositions of the disclosure. In particular, the disclosure provides methods and processes for using any one of the polypeptides or compositions provided herein for hydrolyzing a biomass material. Further, the disclosure provides methods of using any one of the polypeptides or compositions provided herein for reducing the viscosity of a biomass mixture (*e.g.*, a biomass mixture containing biomass substrate and enzyme during saccharification process). In some aspects, there are provided methods of hydrolyzing a biomass material comprising contacting the biomass material with a non-naturally occurring composition comprising a polypeptide having GH61/endoglucanase activity. In some aspects, the polypeptide is in an amount sufficient to hydrolyze the biomass material.

[0301] The term “biomass,” as used herein, refers to any composition comprising cellulose and/or hemicellulose (including lignin in lignocellulosic biomass materials). As used herein, biomass includes, without limitation, seeds, grains, tubers, plant waste or byproducts of food processing or industrial processing (*e.g.*, stalks), corn (including, *e.g.*, cobs, stover, and the like), grasses (including, *e.g.*, Indian grass, such as *Sorghastrum nutans*; or, switchgrass, *e.g.*, *Panicum* species, such as *Panicum virgatum*), perennial canes (*e.g.*, giant reeds), wood (including, *e.g.*, wood chips, processing waste), paper (including paper waste), pulp, and recycled paper (including, *e.g.*, newspaper, printer paper, and the like). Other biomass materials include, without limitation, potatoes, soybean (*e.g.*, rapeseed), barley, rye, oats, wheat, beets, and sugar cane bagasse. Suitable lignocellulosic biomass materials include, without limitation, seeds, grains, tubers, plant waste or byproducts of food processing or industrial processing (*e.g.*, stalks), corn (including, *e.g.*, cobs, stover, and the like), grasses (*e.g.*, Indian grass, such as *Sorghastrum nutans*; or, switchgrass, *e.g.*, *Panicum* species, such as *Panicum virgatum*), perennial canes, *e.g.*, giant reeds, wood (including, *e.g.*, wood chips, processing waste), paper, pulp, recycled paper (*e.g.*, newspaper), wood pulp, or sawdust. Examples of grasses include, without limitation, Indian grass or switchgrass. Examples of reeds include, without limitation, certain perennial canes such as giant reeds. Examples of paper waste include, without limitation, discarded or used photocopy paper, computer printer paper, notebook paper, notepad paper, typewriter paper, newspapers, magazines, cardboard and paper-based packaging materials.

[0302] The saccharified biomass can be made into a number of bio-based products, *via* processes such as, *e.g.*, microbial fermentation and/or chemical synthesis. As used herein, “microbial fermentation” refers to a process of growing and harvesting fermenting microorganisms under suitable conditions. The fermenting microorganism can be any microorganism suitable for use in a desired fermentation process for the production of bio-based products. Suitable fermenting microorganisms include, without limitation, filamentous fungi, yeast, and bacteria. The saccharified biomass can, *e.g.*, be made it into a fuel (*e.g.*, a biofuel such as a bioethanol, biobutanol, biomethanol, a biopropanol, a biodiesel, a jet fuel, or the like) *via* fermentation and/or chemical synthesis. The saccharified biomass can, *e.g.*, also be made into a commodity chemical (*e.g.*, ascorbic acid, isoprene, 1,3-propanediol), lipids, amino acids, proteins, and enzymes, *via* fermentation and/or chemical synthesis.

[0303] Biomass material may include cellulose, hemicellulose, or a mixture thereof. For example, a biomass material may include glucan and/or xylan.

[0304] In some aspects, there are provided methods of reducing the viscosity of a biomass mixture comprising contacting the biomass mixture with non-naturally occurring composition comprising a polypeptide having GH61/endoglucanase activity. The polypeptide is in an amount sufficient to reduce the viscosity. The biomass mixture may comprise biomass material (*e.g.*, pretreated biomass material). The biomass mixture may comprise an enzyme composition such as any of the enzyme compositions provided herein or a mixture thereof.

[0305] In some aspects, any of the polypeptides, compositions provided herein may be used to hydrolyze substrate such as a biomass material or reduce the viscosity of a substrate-enzyme mixture during saccharification process. The substrate may be a biomass material. The substrate may be isolated cellulose or isolated hemicellulose. The substrate may be glucan and/or xylan. In some aspects, the biomass material is pretreated biomass material.

Pretreatment of biomass material

[0306] Prior to saccharification, a biomass material is preferably subject to one or more pretreatment step(s) in order to render xylan, hemicellulose, cellulose and/or lignin material more accessible or susceptible to enzymes and thus more amenable to hydrolysis by the enzyme(s) and/or enzyme blends/compositions of the disclosure.

[0307] Pretreatment may include chemical, physical, and biological pretreatment. For example, physical pretreatment techniques can include without limitation various types of milling, crushing, steaming/steam explosion, irradiation and hydrothermolysis. Chemical

pretreatment techniques can include without limitation dilute acid, alkaline, organic solvent, ammonia, sulfur dioxide, carbon dioxide, and pH-controlled hydrothermolysis. Biological pretreatment techniques can include without limitation applying lignin-solubilizing microorganisms. The pretreatment can occur from several minutes to several hours, such as from about 1 hour to about 120.

[0308] In some aspects, any of the methods or processes provided herein may further comprise pretreating the biomass material, such as pretreating the biomass with acid or base. The acid or base may be ammonia, sodium hydroxide, or phosphoric acid. The method may further comprise pretreating the biomass material with ammonia. The pretreatment may be steam explosion, pulping, grinding, acid hydrolysis, or combinations thereof.

[0309] In one embodiment, the pretreatment may be by elevated temperature and the addition of either of dilute acid, concentrated acid or dilute alkali solution. The pretreatment solution can added for a time sufficient to at least partially hydrolyze the hemicellulose components and then neutralized

[0310] In an exemplary embodiment, the pretreatment entails subjecting biomass material to a catalyst comprising a dilute solution of a strong acid and a metal salt in a reactor. The biomass material can, *e.g.*, be a raw material or a dried material. This pretreatment can lower the activation energy, or the temperature, of cellulose hydrolysis, ultimately allowing higher yields of fermentable sugars. See, *e.g.*, U.S. Patent Nos. 6,660,506; 6,423,145.

[0311] Another exemplary pretreatment method entails hydrolyzing biomass by subjecting the biomass material to a first hydrolysis step in an aqueous medium at a temperature and a pressure chosen to effectuate primarily depolymerization of hemicellulose without achieving significant depolymerization of cellulose into glucose. This step yields a slurry in which the liquid aqueous phase contains dissolved monosaccharides resulting from depolymerization of hemicellulose, and a solid phase containing cellulose and lignin. The slurry is then subject to a second hydrolysis step under conditions that allow a major portion of the cellulose to be depolymerized, yielding a liquid aqueous phase containing dissolved/soluble depolymerization products of cellulose. See, *e.g.*, U.S. Patent No. 5,536,325.

[0312] A further example of method involves processing a biomass material by one or more stages of dilute acid hydrolysis using about 0.4% to about 2% of a strong acid; followed by treating the unreacted solid lignocellulosic component of the acid hydrolyzed material with alkaline delignification. See, *e.g.*, U.S. Patent No. 6,409,841.

[0313] Another example of pretreatment method comprises prehydrolyzing biomass (*e.g.*, lignocellulosic materials) in a prehydrolysis reactor; adding an acidic liquid to the solid lignocellulosic material to make a mixture; heating the mixture to reaction temperature; maintaining reaction temperature for a period of time sufficient to fractionate the lingo-cellulosic material into a solubilized portion containing at least about 20% of the lignin from the lignocellulosic material, and a solid fraction containing cellulose; separating the solubilized portion from the solid fraction, and removing the solubilized portion while at or near reaction temperature; and recovering the solubilized portion. The cellulose in the solid fraction is rendered more amenable to enzymatic digestion. See, *e.g.*, U.S. Patent 5,705,369.

[0314] Further pretreatment methods can involve the use of hydrogen peroxide H₂O₂. See Gould, 1984, Biotech, and Bioengr. 26:46-52.

[0315] Pretreatment can also comprise contacting a biomass material with stoichiometric amounts of sodium hydroxide and ammonium hydroxide at a very low concentration. See Teixeira *et al.*, 1999, Appl. Biochem. and Biotech. 77-79:19-34. Pretreatment can also comprise contacting a lignocellulose with a chemical (*e.g.*, a base, such as sodium carbonate or potassium hydroxide) at a pH of about 9 to about 14 at moderate temperature, pressure, and pH. See PCT Publication WO2004/081185.

[0316] Ammonia may be used in a pretreatment method. Such a pretreatment method comprises subjecting a biomass material to low ammonia concentration under conditions of high solids. See, *e.g.*, U.S. Patent Publication 20070031918, PCT publication WO 06110901. *Saccharification process and viscosity reduction*

[0317] The present disclosure provides methods of reducing the viscosity of a biomass mixture comprising contacting the biomass mixture with a composition (*e.g.*, a non-naturally occurring composition) comprising a polypeptide having glycosyl hydrolase family 61 (“GH61”) endoglucanase activity in an amount sufficient to reduce the viscosity of the biomass mixture. In some aspects, the biomass mixture comprises a biomass material, fermentable sugar(s), whole cellulase, a composition comprising a polypeptide having cellulase activity, and/or a polypeptide having hemicellulase activity. In some aspects, the viscosity is reduced by at least about 5%, (*e.g.*, at least about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%) compared to the viscosity of a biomass mixture in the absence of a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof). In some aspects of any of the methods described herein,

the biomass material comprises hemicellulose, cellulose, or a mixture thereof. In some aspects, the biomass material comprises glucan, xylan and/or lignin.

[0318] The methods and processes provided herein may be performed under various conditions. For example, any of the methods provided herein may be performed at a pH in the range of pH of about 3.5 to about 7.0, for example, pH of about 4.0 to about 6.5, pH of about 4.4 to about 6.0, pH of about 4.8 to about 5.6, or about 4.5 to about 5.5. The saccharification mixture containing biomass material may be adjusted to the desired pH using base or acid (such as sulfuric acid) according to any of the methods known to one of ordinary skill in the art. For example, the pretreated biomass material may be added with base or acid (such as sulfuric acid) to achieve the desired pH for saccharification. Any of the methods for hydrolyzing a biomass material or reducing the viscosity of the biomass mixture may be conducted at a pH of about 4.8 to about 5.6 (*e.g.*, pH of about any of 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, or 5.6). In some aspects, the method further comprises adjusting the pH of the biomass mixture to a pH of about 4.0 to about 6.5 (*e.g.*, pH of about 4.5 to about 5.5).

[0319] The methods and processes provided herein may be performed for any length of time, *e.g.*, 1 hour, 2 hours, 4 hours, 8 hours, 12 hours, 18 hours, 24 hours, 2 days, 3 days, 4 days, 5 days, 6 days, 7 days, 8 days, 10 days, 14 days, 3 weeks, or 4 weeks. After any of the saccharification time described herein, the amount of fermentable sugar(s) is increased and/or the viscosity of the saccharification mixture is reduced. In some aspects, the method is performed for about 2 hours to about 7 days (*e.g.*, about 4 hours to about 6 days, about 8 hours to about 5 days, or about 8 hours to about 3 days).

[0320] A composition (*e.g.*, a non-naturally occurring composition) comprising polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof) may be added after the biomass material is pretreated. A composition (*e.g.*, a non-naturally occurring composition) comprising polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof) may be added to the biomass material before or after another enzyme composition (such as an enzyme composition comprising hemicellulose, cellulase, or whole cellulase) is added to the biomass material. A composition (*e.g.*, a non-naturally occurring composition) comprising polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof) may be added to the biomass mixture containing (a) biomass material and/or fermentable sugars and (b) enzyme (such as hemicellulase or cellulase including whole cellulase). In some aspects, a

composition (*e.g.*, a non-naturally occurring composition) comprising polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof) is added to the biomass mixture, wherein the biomass material has been hydrolyzed for a period of time (such as about any of 5 minutes, 10 minutes, 30 minutes, 1 hour, 2 hours, 4 hours, 8 hours, 12 hours, 18 hours, 24 hours, 2 days, 3 days, 4 days, or 5 days).

[0321] A composition (*e.g.*, a non-naturally occurring composition) comprising isolated polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof) may be added to biomass material during saccharification. A composition (*e.g.*, a non-naturally occurring composition) comprising whole cellulase may be added to biomass material during saccharification, where the whole cellulase comprises a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof).

[0322] A biomass material used in any one of the methods may be in liquid form, solid form, or a mixture thereof. A biomass material used in any one of the methods may be wet form, dry form, a material having various degree of moisture, or a mixture thereof. A biomass material used in any one of the methods may be in a dry solid form (such as a dry solid form as a starting material). The biomass material may be processed into any of the following forms: wet form, dry form, solid form, liquid form, or a mixture thereof according to any method known to one skilled in the art.

[0323] A biomass material used in any of the methods may be present in the saccharification mixture in an amount of at least about any of 0.5 wt.%, 1 wt.%, 5 wt.%, 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, or 60 wt.% of total weight of hydrolysis mixture or saccharification mixture, wherein the amount of the biomass material refers to the weight amount of the biomass material in its solid state (or the biomass material in its dry state, its dry solid state, its natural state, or its unprocessed state). The biomass material may also be in an amount of about 0.5 wt.% to about 55 wt.%, 1 wt.% to about 40 wt.%, 5 wt.% to about 60 wt.%, about 10 wt.% to about 55 wt.%, about 10 wt.% to about 50 wt.%, about 15 wt.% to about 50 wt.%, about 15 wt.% to about 40 wt.%, about 15 wt.% to about 35 wt.%, about 15 wt.% to about 30 wt.%, about 20 wt.% to about 35 wt.%, or about 20 wt.% to about 30 wt.% of a hydrolyzing mixture containing biomass material, wherein the amount of the biomass material refers to the weight amount of the biomass material in its solid state (or the biomass material in its dry state, its dry solid state, its natural state, or its unprocessed state). A biomass material used in any of

the methods may be present in the saccharification mixture in an amount of about any of 0.5 wt.%, 1 wt.%, 5 wt.%, 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 45 wt.%, 50 wt.%, 55 wt.%, or 60 wt.% of total weight of hydrolysis mixture or saccharification mixture, wherein the amount of the biomass material refers to the weight amount of the biomass material in its solid state (or the biomass material in its dry state, its dry solid state, its natural state, or its unprocessed state).

[0324] The hydrolysis mixture or saccharification mixture includes biomass material, enzyme(s) (*e.g.*, any one of polypeptides provided herein), enzyme composition (*e.g.*, any one of the compositions provided herein), and/or other components such as components necessary for saccharification.

[0325] Any of the compositions provided herein may be used in the methods described herein such as any one of the compositions provided above in the "Exemplary compositions" section. The amount of any of the compositions described herein used in any one of the methods provided herein may be in the range of about 0.05 mg to about 50 mg, about 0.1 mg to about 40 mg, about 0.2 mg to about 30 mg, about 0.5 mg to about 25 mg, about 1 mg to about 25 mg, about 2 mg to about 25 mg, about 5 mg to about 25 mg, or about 10 mg to about 25 mg protein per gram of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose contained in the biomass material. A non-naturally occurring composition comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof) used in any one of the methods for hydrolyzing a biomass material and/or methods for reducing the viscosity of the biomass mixture may be in an amount of about 0.05 mg to about 50 mg, about 0.1 mg to about 40 mg, about 0.2 mg to about 30 mg, about 0.5 mg to about 25 mg, about 1 mg to about 25 mg, about 2 mg to about 25 mg, about 5 mg to about 25 mg, or about 10 mg to about 25 mg protein per gram of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose contained in the substrate such as biomass material.

[0326] In some aspects, a non-naturally occurring composition comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof) used in any of the methods for hydrolyzing a biomass material and/or methods for reducing the viscosity of the biomass mixture is in an amount of at least about any of 0.05 mg, 0.1 mg, 0.2 mg, 0.5 mg, 1 mg, 2 mg, 5 mg, 7.5 mg, 10 mg, 12 mg, 14 mg, 15 mg, 16 mg, 17.5 mg, 18 mg, 20 mg, 22.5 mg, 25 mg, 27.5 mg, 30 mg, 35 mg, 40 mg, 45 mg, or 50 mg protein per

gram of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose contained in the substrate such as biomass material. In some aspects, a non-naturally occurring composition comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof) used in any of the methods for hydrolyzing a biomass material

5 and/or methods for reducing the viscosity of the biomass mixture is in an amount of no more than about any of 0.1 mg, 0.2 mg, 0.5 mg, 1 mg, 2 mg, 5 mg, 7.5 mg, 10 mg, 12 mg, 14 mg, 15 mg, 16 mg, 17.5 mg, 18 mg, 20 mg, 22.5 mg, 25 mg, 27.5 g mg, 30 mg, 35 mg, 40 mg, 45 mg, 50 mg, 55 mg, 60 mg, 65 mg, 75 mg, or 100 mg protein per gram of cellulose,

hemicellulose, or a mixture of cellulose and hemicellulose contained in the substrate such as

10 biomass material. In some aspects, a non-naturally occurring composition comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof) used in any of the methods for hydrolyzing a biomass material and/or methods for reducing the viscosity of the biomass mixture is in an amount of about any of 0.05 mg, 0.1 mg, 0.2 mg, 0.5 mg, 1 mg, 2 mg, 5 mg, 7.5 mg, 10 mg, 12 mg, 14 mg, 15 mg, 16 mg, 17.5 mg, 18 mg, 20 mg, 22.5 mg, 25 mg, 27.5 g mg, 30 mg, 35 mg, 40 mg, 45 mg, or 50 mg protein per gram of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose

15 contained in the substrate such as biomass material. The amount of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose contained in the substrate such as biomass material may be calculated using any methods known to one skilled in the art. The biomass material may comprise glucan, xylan, and/or lignin.

[0327] In some aspects of any of the methods described herein, the amount of the composition comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is about 0.1 mg to about 50 mg protein (*e.g.*, about 0.2 mg to about 40 mg protein, about 0.5 mg to about 30 mg protein, about 1 mg to about 20 mg protein, or
25 about 5 mg to about 15 mg protein) per gram of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose contained in the biomass material. The protein amount described herein refers to the weight of total protein in the composition. The proteins include a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and may also include other enzymes such as cellulase polypeptide(s) and/or hemicellulase
30 polypeptide(s) in the composition.

[0328] In some aspects of any of the methods described herein, the amount of the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is

about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg protein, about 0.5 mg to about 10 mg protein, or about 1 mg to about 5 mg protein) per gram of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose contained in the biomass material.

[0329] In some aspects of any of the methods described herein, the composition comprises

5 a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one polypeptide having endoglucanase activity (*e.g.*, *T. reesei* Eg1, *T. reesei* Eg2, and/or a variant thereof), wherein the total amount of the polypeptides having endoglucanase activity is about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg protein, about 0.5 mg to about 10 mg protein, or about 1 mg to about 5 mg protein) per gram of cellulose,

10 hemicellulose, or a mixture of cellulose and hemicellulose contained in the biomass material.

[0330] In some aspects, the composition comprises a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one polypeptide having cellobiohydrolase activity (*e.g.*, *T. reesei* CBH1, *T. reesei* CBH2, and/or a variant thereof), wherein the amount of the polypeptide(s) having cellobiohydrolase activity is about

15 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg protein, about 0.5 mg to about 10 mg protein, or about 1 mg to about 5 mg protein) per gram of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose contained in the biomass material.

[0331] In some aspects of any of the methods described herein, the composition comprises a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof)

20 and at least one polypeptide having β -glucosidase activity (*e.g.*, an Fv3C, a Pa3D, an Fv3G, an Fv3D, a Tr3A, a Tr3B, a Te3A, an An3A, an Fo3A, a Gz3A, an Nh3A, a Vd3A, a Pa3G, a Tn3B, or a variant thereof), wherein the amount of the polypeptide(s) having β -glucosidase activity is about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg protein, about 0.5 mg to about 10 mg protein, or about 0.5 mg to about 5 mg protein) per gram of cellulose,

25 hemicellulose, or a mixture of cellulose and hemicellulose contained in the biomass material.

[0332] In some aspects, the composition comprises a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one polypeptide having xylanase activity (*e.g.*, *T. reesei* Xyn3, *T. reesei* Xyn2, an AfuXyn2, an AfuXyn5, or a variant thereof), wherein the amount of the polypeptide(s) having xylanase activity is about

30 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg protein, about 0.5 mg to about 10 mg protein, or about 0.5 mg to about 5 mg protein) per gram of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose contained in the biomass material.

[0333] In some aspects, the composition comprises a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one polypeptide having β -xylosidase activity (*e.g.*, Fv3A, Fv43A, a Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fo43A, an Fv43B, a Pa51A, a Gz43A, a *T. reesei* Bx11, or a variant thereof), wherein the amount of the polypeptide(s) having β -xylosidase activity is about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg protein, about 0.5 mg to about 10 mg protein, or about 0.5 mg to about 5 mg protein) per gram of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose contained in the biomass material.

[0334] In some aspects, the composition comprises a polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one polypeptide having L- α -arabinofuranosidase activity (*e.g.*, an Af43A, an Fv43B, a Pf51A, a Pa51A, an Fv51A, or a variant thereof), wherein the amount of the polypeptide(s) having L- α -arabinofuranosidase activity is about 0.2 mg to about 30 mg (*e.g.*, about 0.2 mg to about 20 mg protein, about 0.5 mg to about 10 mg protein, or about 0.5 mg to about 5 mg protein) per gram of cellulose, hemicellulose, or a mixture of cellulose and hemicellulose contained in the biomass material.

[0335] In any one of the methods provided herein, the viscosity of the biomass mixture may be reduced by at least about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% compared to the viscosity of the biomass mixture in the absence of an enzyme composition provided herein. For example, there are provided methods of reducing the viscosity of a biomass mixture comprising contacting the biomass mixture with a non-naturally occurring composition comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof), wherein the viscosity is reduced by at least about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% compared to the viscosity of the biomass mixture in the absence of a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof). In some aspects, the viscosity is reduced by about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% compared to the viscosity of the biomass mixture in the absence of a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof). The reduction of viscosity described herein is seen after a certain period of saccharification. For example, the reduction

of viscosity is seen after 30 minutes, 1 hour, 2 hours, 4 hours, 8 hours, 12 hours, 18 hours, 24 hours, 2 days, 3 days, 4 days, or 5 days saccharification. Methods of measuring viscosity are known in the art. For example, viscosity may be measured by human eyes, or be measured by a viscometer such as Brookfield viscometer (Brookfield Engineering, Inc). For example, viscosity of saccharification reaction mixture can be measured using a viscosity meter with ammonia-pretreated corncob as substrates. A viscosity meter can measure the resistance (torque) it takes to turn a spindle at a constant rate in the slurry.

[0336] The methods provided herein may be conducted at a temperature that is suitable for saccharification. For example, any one of the methods described herein may be performed at about 20°C to about 75°C, about 25°C to about 70°C, about 30°C to about 65°C, about 35°C to about 60°C, about 37°C to about 60°C, about 40°C to about 60°C, about 40°C to about 55°C, about 40°C to about 50°C, or about 45°C to about 50°C. In some aspects, any one of the methods described herein may be performed at about 20°C, about 25°C, about 30°C, about 35°C, about 37°C, about 40°C, about 45°C, about 48°C, about 50°C, about 55°C, about 60°C, about 65°C, about 70°C, or about 75°C.

[0337] In some aspects of any of the methods described herein, the method comprises producing fermentable sugar(s), wherein the amount of the fermentable sugar(s) is increased by at least about 5% (e.g., at least about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, or 90%) compared to the amount of the fermentable sugar(s) produced in the absence of a polypeptide having GH61/endoglucanase activity (e.g., *T. reesei* Eg4 or a variant thereof).

[0338] Also provided herein are methods of increasing the amount of fermentable sugar(s) (and/or increasing the conversion from a biomass material to fermentable sugar(s) such as glucan conversion) by using a composition (e.g., a non-naturally occurring composition) comprising a polypeptide having GH61/endoglucanase activity (e.g., EG IV such as *T. reesei* Eg4 or a variant thereof) during hydrolysis of biomass material. There are various fermentable sugars produced from hydrolysis of biomass material, including but are not limited to, glucose, xylose, and/or cellobiose. In some aspects, the amount of fermentable sugar(s) produced from hydrolysis of biomass material may be increased by at least about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% compared to the amount of fermentable sugar(s) in the absence of an enzyme composition provided herein. For example, there are provided methods of

increasing the amount of fermentable sugar(s) comprising contacting the biomass material with a non-naturally occurring composition comprising a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof) (to start or further a saccharification process), wherein the amount of fermentable sugar(s) from saccharification is increased by at least about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% compared to the amount of fermentable sugar(s) from saccharification in the absence of a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof). In some aspects, the amount of fermentable sugar(s) from saccharification is increased by about any of 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% compared to the amount of fermentable sugar(s) from saccharification in the absence of a polypeptide having GH61/endoglucanase activity (*e.g.*, EG IV such as *T. reesei* Eg4 or a variant thereof). The increase in amount of fermentable sugar(s) produced from hydrolysis of biomass material described herein is seen after a certain period of saccharification. For example, the increase in amount of fermentable sugar(s) is seen after 30 minutes, 1 hour, 2 hours, 4 hours, 8 hours, 12 hours, 18 hours, 24 hours, 2 days, 3 days, 4 days, or 5 days saccharification. Methods of measuring amount of fermentable sugar(s) and/or glucan conversion are known to a person skilled in the art.

[0339] The reduction in viscosity of saccharification mixture may correlate with improved yield of desirable fermentable sugars.

[0340] In some aspects, the method further comprises the step of contacting the biomass material with a composition comprising whole cellulase. In some aspects, the step of further contacting the biomass material with a composition comprising whole cellulase is performed before, after, or concurrently with contacting the biomass material with composition comprising a polypeptide having glycosyl hydrolase family 61 ("GH61") endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof).

[0341] In some aspects of any of the methods described herein, the method comprises the step of further contacting the biomass material with a composition comprising a polypeptide having cellulase activity and/or a polypeptide having hemicellulase activity. In some aspects, the step of further contacting the biomass material with a composition comprising a polypeptide having cellulase activity and/or a polypeptide having hemicellulase activity is performed before, after, or concurrently with contacting the biomass material with

composition comprising a polypeptide having glycosyl hydrolase family 61 ("GH61") endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof).

[0342] In some aspects, the composition comprises the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and further comprises at least one cellulase polypeptide and/or at least one hemicellulase polypeptide, wherein the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one cellulase polypeptide and/or at least one hemicellulase polypeptide are mixed together before contacting the biomass material with a composition comprising the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof).

[0343] In some aspects, the composition comprises the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and further comprises at least one cellulase polypeptide and/or at least one hemicellulase polypeptide, wherein the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) and at least one cellulase polypeptide and/or at least one hemicellulase polypeptide are added to the biomass material at different times (*e.g.*, the polypeptide having GH61/endoglucanase activity (*e.g.*, *T. reesei* Eg4 or a variant thereof) is added before or after at least one cellulase polypeptide and/or at least one hemicellulase polypeptide is added to the biomass material).

[0344] Enhanced cellulose conversion may be achieved at higher temperatures using the CBH polypeptides described in, for example, any one of the following US Patent

Publications US20050054039, US20050037459, US20060205042, US20050048619A1 and US20060218671. Methods of overexpressing β -glucosidase are known in the art. See, *e.g.*, U.S. Patent 6,022,725. See also, *e.g.*, US Patent Publication 20050214920.

[0345] The methods of the present disclosure can be used in the production of monosaccharides, disaccharides, and polysaccharides as chemical, fermentation feedstocks for microorganism, and inducers for the production of proteins, organic products, chemicals and fuels, plastics, and other products or intermediates. In particular, the value of processing residues (dried distillers grain, spent grains from brewing, sugarcane bagasse, etc.) can be increased by partial or complete solubilization of cellulose or hemicellulose. In addition to ethanol, chemicals that can be produced from cellulose and hemicellulose include, acetone, acetate, glycine, lysine, organic acids (*e.g.*, lactic acid), 1,3-propanediol, butanediol, glycerol, ethylene glycol, furfural, polyhydroxyalkanoates, cis, cis-muconic acid, animal feed and xylose.

Business Methods

[0346] The cellulase and/or hemicellulase compositions of the disclosure can be further used in industrial and/or commercial settings. Accordingly a method or a method of manufacturing, marketing, or otherwise commercializing the instant non-naturally occurring cellulase and/or hemicellulase compositions is also contemplated.

[0347] In a specific embodiment, the non-naturally occurring cellulase and/or hemicellulase compositions of the invention, for example, comprising one or more of the GH61 endoglucanases or variants thereof as described herein, can be supplied or sold to certain ethanol (bioethanol) refineries or other bio-chemical or bio-material manufacturers. In a first example, the non-naturally occurring cellulase and/or hemicellulase compositions can be manufactured in an enzyme manufacturing facility that is specialized in manufacturing enzymes at an industrial scale. The non-naturally occurring cellulase and/or hemicellulase compositions can then be packaged or sold to customers of the enzyme manufacturer. This operational strategy is termed the “merchant enzyme supply model” herein.

[0348] In another operational strategy, the non-naturally occurring cellulase and hemicellulase compositions of the invention can be produced in a state of the art enzyme production system that is built by the enzyme manufacturer at a site that is located at or in the vicinity of the bioethanol refineries or the bio-chemical/biomaterial manufacturers (“on-site”). In some embodiments, an enzyme supply agreement is executed by the enzyme manufacturer and the bioethanol refinery or the bio-chemical/biomaterial manufacturer. The enzyme manufacturer designs, controls and operates the enzyme production system on site, utilizing the host cell, expression, and production methods as described herein to produce the non-naturally-occurring cellulase and/or hemicellulase compositions. In certain embodiments, suitable biomass, preferably subject to appropriate pretreatments as described herein, can be hydrolyzed using the saccharification methods and the enzymes and/or enzyme compositions herein at or near the bioethanol refineries or the bio-chemical/biomaterial manufacturing facilities. The resulting fermentable sugars can then be subject to fermentation at the same facilities or at facilities in the vicinity. This operational strategy is termed the “on-site biorefinery model” herein.

[0349] The on-site biorefinery model provides certain advantages over the merchant enzyme supply model, including, e.g., the provision of a self-sufficient operation, allowing minimal reliance on enzyme supply from merchant enzyme suppliers. This in turn allows the

bioethanol refineries or the bio-chemical/biomaterial manufacturers to better control enzyme supply based on real-time or nearly real-time demand. In certain embodiments, it is contemplated that an on-site enzyme production facility can be shared between two or among two or more bioethanol refineries and/or the bio-chemical/biomaterial manufacturers who are
5 located near to each other, reducing the cost of transporting and storing enzymes. Moreover, this allows more immediate “drop-in” technology improvements at the enzyme production facility on-site, reducing the time lag between the improvements of enzyme compositions to a higher yield of fermentable sugars and ultimately, bioethanol or biochemicals.

[0350] The on-site biorefinery model has more general applicability in the industrial

10 production and commercialization of bioethanols and biochemicals, in that it can be used to manufacture, supply, and produce not only the cellulase and non-naturally occurring hemicellulase compositions of the present disclosure but also those enzymes and enzyme compositions that process starch (*e.g.*, corn) to allow for more efficient and effective direct conversion of starch to bioethanol or bio-chemicals. The starch-processing enzymes can, in
15 certain embodiments, be produced in the on-site biorefinery, then quickly and easily integrated into the bioethanol refinery or the biochemical/biomaterial manufacturing facility in order to produce bioethanol.

[0096] Thus in certain aspects, the invention also pertains to certain business method of applying the enzymes (*e.g.*, certain GH61 endoglucanases and variants thereof), cells,

20 compositions (*e.g.*, comprising a suitable GH61 endoglucanase or a variant thereof), and processes herein in the manufacturing and marketing of certain bioethanol, biofuel, biochemicals or other biomaterials. In some embodiments, the invention pertains to the application of such enzymes, cells, compositions and processes in an on-site biorefinery model. In other embodiments, the invention pertains to the application of such enzymes,
25 cells, compositions and processes in a merchant enzyme supply model.

[0097] Relatedly, the disclosure provides the use of the enzymes and/or the enzyme compositions of the invention in a commercial setting. For example, the enzymes and/or enzyme compositions of the disclosure can be sold in a suitable market place together with instructions for typical or preferred methods of using the enzymes and/or compositions.

30 Accordingly the enzymes and/or enzyme compositions of the disclosure can be used or commercialized within a merchant enzyme supplier model, where the enzymes and/or enzyme compositions of the disclosure are sold to a manufacturer of bioethanol, a fuel

refinery, or a biochemical or biomaterials manufacturer in the business of producing fuels or bio-products. In some aspects, the enzyme and/or enzyme composition of the disclosure can be marketed or commercialized using an on-site bio-refinery model, wherein the enzyme and/or enzyme composition is produced or prepared in a facility at or near to a fuel refinery or biochemical/biomaterial manufacturer's facility, and the enzyme and/or enzyme composition of the invention is tailored to the specific needs of the fuel refinery or biochemical/biomaterial manufacturer on a real-time basis. Moreover, the disclosure relates to providing these manufacturers with technical support and/or instructions for using the enzymes and/or enzyme compositions such that the desired bio-product (*e.g.*, biofuel, biochemicals, bio-materials, etc) can be manufactured and marketed.

[0351] The following are examples of the methods and compositions of the invention. It is understood that various other embodiments may be practiced, given the general description provided above.

EXAMPLES

Example 1: Assays/Methods

[0352] The following assays/methods were generally used in the Examples described below. Any deviations from the protocols provided below are indicated in specific Examples.

A. Pretreatment of biomass substrates

[0353] Corncob, corn stover and switch grass were pretreated prior to enzymatic hydrolysis according to the methods and processing ranges described in International Patent Publication WO06110901A (unless otherwise noted). These references for pretreatment are also included in the disclosures of US Patent Application Publications 20070031918-A1, 20070031919-A1, 20070031953-A1, and/or 20070037259-A1.

[0354] Ammonia fiber explosion treated (AFEX) corn stover was obtained from Michigan Biotechnology Institute International (MBI). The composition of the corn stover was determined by MBI (Teymouri, F et al. Applied Biochemistry and Biotechnology, 2004, 113:951-963) using the National Renewable Energy Laboratory (NREL) procedure, NREL LAP-002. NREL procedures are available at: http://www.nrel.gov/biomass/analytical_procedures.html.

[0355] The FPP pulp and paper substrates were obtained from SMURFIT KAPPA CELLULOSE DU PIN, France.

[0356] Steam Expanded Sugar-cane Bagasse (SEB) was obtained from SunOpta (Glasser, WG et al. Biomass and Bioenergy 1998, 14(3): 219-235; Jollez, P et al. Advances in thermochemical biomass conversion, 1994, 2:1659-1669).

B. Compositional analysis of biomass

5 [0357] The 2-step acid hydrolysis method described in Determination of structural carbohydrates and lignin in the biomass (National Renewable Energy Laboratory, Golden, CO 2008 <http://www.nrel.gov/biomass/pdfs/42618.pdf>) was used to measure the composition of biomass substrates. Using this method, enzymatic hydrolysis results were reported herein in terms of percent conversion with respect to the theoretical yield from the starting glucan
10 and xylan content of the substrate.

C. Total protein assay

[0358] The BCA protein assay is a colorimetric assay that measures protein concentration with a spectrophotometer. The BCA Protein Assay Kit (Pierce Chemical, Product #23227) was used according to the manufacturer's suggestion. Enzyme dilutions were prepared in test
15 tubes using 50 mM sodium acetate pH 5 buffer. Diluted enzyme solution (0.1 mL) was added to 2 mL Eppendorf centrifuge tubes containing 1 mL 15% trichloroacetic acid (TCA). The tubes were vortexed and placed in an ice bath for 10 min. The samples were then centrifuged at 14,000 rpm for 6 min. The supernatant was poured out, the pellet was resuspended in 1 mL 0.1 N NaOH, and the tubes vortexed until the pellet dissolved. BSA standard solutions were
20 prepared from a stock solution of 2 mg/mL. BCA working solution was prepared by mixing 0.5 mL Reagent B with 25 mL Reagent A. 0.1 mL of the enzyme resuspended sample was added to 3 Eppendorf centrifuge tubes. Two (2) mL Pierce BCA working solution was added to each sample and BSA standard Eppendorf tubes. All tubes were incubated in a 37°C waterbath for 30 min. The samples were then cooled to room temperature (15 min) and the
25 absorbance measured at 562 nm in a spectrophotometer.

[0359] Average values for the protein absorbance for each standard were calculated. The average protein standard was plotted, absorbance on x-axis and concentration (mg/mL) on the y-axis. The points were fit to a linear equation:

$$y=mx +b$$

30 [0360] The raw concentration of the enzyme samples was calculated by substituting the absorbance for the x- value. The total protein concentration was calculated by multiplying with the dilution factor.

[0361] The total protein of purified samples was determined by A280 (Pace, CN, et al. *Protein Science*, 1995, 4:2411-2423).

[0362] The total protein content of fermentation products was sometimes measured as total nitrogen by combustion, capture and measurement of released nitrogen, either by Kjeldahl (rtech laboratories, www.rtechlabs.com) or in-house by the DUMAS method (TruSpec CN, www.leco.com) (Sader, A.P.O. et al., *Archives of Veterinary Science*, 2004, 9(2):73-79). For complex protein-containing samples, e.g. fermentation broths, an average 16% N content, and the conversion factor of 6.25 for nitrogen to protein was used. In some cases, total precipitable protein was measured to remove interfering non-protein nitrogen. A 12.5% final TCA concentration was used and the protein-containing TCA pellet was resuspended in 0.1 M NaOH.

[0363] In some cases, Coomassie Plus- the Better Bradford Assay (Thermo Scientific, Rockford, IL product #23238) was used according to manufacturer recommendation. In other cases, total protein was measured using the Biuret method as modified by Weichselbaum and Gornall using Bovine Serum Albumin as a calibrator (Weichselbaum, T. *Amer. J. Clin. Path.* 1960,16:40; Gornall, A. et al. *J. Biol. Chem.* 1949, 177:752).

D. Glucose determination using ABTS

[0364] The ABTS (2, 2'-azino-bis(3-ethylenethiazoline-6)-sulfonic acid) assay for glucose determination is based on the principle that in the presence of O₂, glucose oxidase catalyzes the oxidation of glucose while producing stoichiometric amounts of hydrogen peroxide (H₂O₂). This reaction is followed by the horse radish peroxidase (HRP) catalyzed oxidation of ABTS which linearly correlates to the concentration of H₂O₂. The emergence of oxidized ABTS is indicated by the evolution of a green color, which is quantified at an OD of 405 nm. A mixture of ABTS powder (Sigma, #A1888-5g 2.74 mg/mL), 0.1 U/mL HRP (100 U/mL, Sigma, #P8375) and 1 U/mL Glucose Oxidase, (OxyGO® HP L5000, 5000 U/mL, Genencor Division, Danisco USA) was prepared in 50 mM Na Acetate Buffer, pH 5.0 and kept in the dark (substrate). Glucose standards (0, 2, 4, 6, 8, 10 nmol) were prepared in 50 mM Na Acetate Buffer, pH 5.0 and 10 µL of each standard was added to a 96-well flat bottom MTP in triplicate. Ten (10) µL of serially diluted samples were also added to the MTP. One hundred (100) µL of ABTS substrate solution was added to each well and the plate was placed on a spectrophotometric plate reader to kinetically read oxidation of ABTS for 5 min at 405 nm.

[0365] Alternately absorbance at 405 nm was measured after 15-30 min of incubation followed by quenching of the reaction with 50 mM Na Acetate Buffer, pH 5.0 containing 2% SDS.

E. Sugar analysis by HPLC

5 [0366] Samples from biomass saccharification were prepared by centrifugation to clear insoluble material, filtration through a 0.22 μ m nylon filter (Spin-X centrifuge tube filter, Corning Incorporated, Corning, NY) and dilution to an appropriate concentration of soluble sugars with distilled water. Monomer sugars were determined on a Shodex Sugar SH-G SH1011, 8x300 mm with a 6x50 mm SH-1011P guard column (www.shodex.net). Solvent
10 was 0.01 N H₂SO₄ run at 0.6 mL/min. Column temperature was 50°C and detection was by refractive index. Alternately, sugars were analyzed using a Biorad Aminex HPX-87H column with a Waters 2410 refractive index detector. The analysis time was 20 min, the injection volume was 20 μ L of diluted sample, the mobile phase was 0.01 N sulfuric acid, 0.2 μ m filtered and degassed, the flow rate was 0.6 mL/min and the column temperature was 60°C.

15 External standards of glucose, xylose and arabinose were run with each sample set.

[0367] Oligomeric sugars were separated by size exclusion chromatography in HPLC using a Tosoh Biosep G2000PW column 7.5 mmx60 cm (www.tosohbioscience.de). The solvent was distilled water at 0.6 mL/min and the column was run at room temperature. Six carbon sugar standards used for size calibration were: stachyose, raffinose, cellobiose and glucose;
20 and 5 carbon sugars were: xylohexose, xylopentose, xylotetrose, xylotriose, xylobiose and xylose. Xylo-oligomers were obtained from Megazyme (www.megazyme.com). Detection was by refractive index and when reported quantitatively results are either as peak area units or relative peak areas by percent.

[0368] Total soluble sugars were determined by hydrolysis of the centrifuged and filter
25 clarified samples described above. The clarified sample was diluted 1 to 1 with 0.8 N H₂SO₄ and the resulting solution was autoclaved in a capped vial for a total cycle time of 1 h at 121°C. Results are reported without correction for loss of monomer sugar during the hydrolysis.

F. Oligomer Preparation from Cob and Enzyme Assays

30 [0369] Oligomers from *T. reesei* Xyn3 hydrolysis of corncobs were prepared by incubating 8 mg *T. reesei* Xyn3 per g Glucan + Xylan with 250 g dry weight of dilute ammonia pretreated corncob in 50 mM pH 5.0 Na Acetate buffer (pH adjusted with 1 N sulfuric acid).

The reaction proceeded for 72 h at 48°C, 180 rpm rotary shaking. The supernatant was centrifuged 9,000 x G, then filtered through 0.22 µm Nalgene filters to recover the soluble sugars. For subsequent enzyme assays, 100 µL aliquots of the *T. reesei* Xyn3 oligomer-containing supernatant were incubated with 1 µg/µL of either *T. reesei* integrated strain H3A, 1 µg/µL of *T. reesei* integrated strain H3A/EG4#27 or water control in Eppendorf tubes at 48°C for 2.5 h. The supernatants were then diluted 4X with ice cold MilliQ water, filtered, and analyzed by HPLC for sugar release from the oligomers.

G. Corncob Saccharification Assay

[0370] For a typical example herein, unless otherwise specifically described with the particular examples, corncob saccharification was performed in a microtiter plate format in accordance with the following procedures. The biomass substrate, *e.g.*, a dilute ammonia pretreated corncob, was diluted in water and pH-adjusted with sulfuric acid to create a pH 5, 7% cellulose slurry that was then used directly without further processing in the assays. Enzyme samples were loaded based on mg total protein per g of cellulose (as determined using conventional compositional analysis methods, such as, for example, using the method described in Example 1A above) in the substrate (*e.g.*, the corncob). The enzymes were then diluted in 50 mM sodium acetate, pH 5.0, to obtain the desired loading concentration. Forty (40) µL of enzyme solution were added to 70 mg of dilute-ammonia pretreated corncob at 7% cellulose per well (equivalent to 4.5% cellulose final per well). The assay plates were covered with aluminum plate sealers, mixed at room temperature and incubated at 50°C, 200 rpm, for 3 days (“3d”). At the end of the incubation period, the saccharification reaction was quenched by adding to each well 100 µL of a 100 mM glycine buffer, pH10.0. The plate was centrifuged for 5 min at 3,000 rpm. Ten (10) µL of the supernatant was then added to 200 µL of MilliQ water in a 96-well HPLC plate and the soluble sugars were measured using HPLC.

Example 2: Construction of an Integrated Expression Strain of *Trichoderma reesei*

[0371] An integrated expression strain of *Trichoderma reesei* was constructed that co-expressed five genes: *T. reesei* β-glucosidase gene *bgl1*, *T. reesei* endoxylanase gene *xyn3*, *F. verticillioides* β-xylosidase gene *fv3A*, *F. verticillioides* β-xylosidase gene *fv43D*, and *F. verticillioides* α-arabinofuranosidase gene *fv51A*.

[0372] The construction of the expression cassettes for these different genes and the transformation of *T. reesei* are described below.

A. Construction of the β-glucosidase expression vector

[0373] The N-terminal portion of the native *T. reesei* β -glucosidase gene *bglI* was codon optimized by DNA 2.0 (Menlo Park, USA). This synthesized portion comprised of the first 447 bases of the coding region. This fragment was PCR amplified using primers SK943 and SK941. The remaining region of the native *bglI* gene was PCR amplified from a genomic

DNA sample extracted from *T. reesei* strain RL-P37 (Sheir-Neiss, G *et al.* Appl. Microbiol. Biotechnol. 1984, 20:46-53), using primer SK940 and SK942. These two PCR fragments of the *bglI* gene were fused together in a fusion PCR reaction, using primers SK943 and SK942: Forward Primer SK943: (5'-CACCATGAGATATAGAACAGCTGCCGCT-3') (SEQ ID NO:121)

Reverse Primer SK941: (5'-CGACCGCCCTGCGGAGTCTTGCCAGTGGTCCCGCGACAG-3') (SEQ ID NO:122)
Forward Primer (SK940): (5'-CTGTGCGGGGACCACTGGGCAAGACTCCGCAGGGCGGTGCG-3') (SEQ ID NO:123)

Reverse Primer (SK942): (5'-CCTACGCTACCGACAGAGTG-3') (SEQ ID NO:124)

[0374] The resulting fusion PCR fragments were cloned into the Gateway® Entry vector pENTR™/D-TOPO®, and transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen) resulting in the intermediate vector, pENTR-TOPO-BglI-(943/942) (FIG. 8A). The nucleotide sequence of the inserted DNA was determined. The pENTR-943/942 vector with the correct *bglI* sequence was recombined with pTrex3g using a LR clonase® reaction protocol outlined by Invitrogen. The LR clonase reaction mixture was transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen), resulting in the final expression vector, pTrex3g 943/942 (FIG. 8B). The vector also contains the *Aspergillus nidulans amdS* gene, encoding acetamidase, as a selectable marker for transformation of *T. reesei*. The expression cassette was amplified by PCR with primers SK745 and SK771 to generate product for transformation of *T. reesei*.

Forward Primer SK771: (5'-GTCTAGACTGGAAACGCAAC -3') (SEQ ID NO:125)

Reverse Primer SK745: (5'-GAGTTGTGAAGTCGGTAATCC -3') (SEQ ID NO:126)

B. Construction of the endoxylanase expression cassette

[0375] The native *T. reesei* endoxylanase gene *xyn3* was PCR amplified from a genomic DNA sample extracted from *T. reesei*, using primers xyn3F-2 and xyn3R-2.

Forward Primer xyn3F-2: (5'-CACCATGAAAGCAAACGTCATCTTGTGCCTCCTGG-3') (SEQ ID NO:127)

Reverse Primer (xyn3R-2): (5'-CTATTGTAAGATGCCAACAATGCTGTTATATGC
CGGCTTGGGG-3') (SEQ ID NO:128)

[0376] The resulting PCR fragments were cloned into the Gateway® Entry vector pENTR™/D-TOPO®, and transformed into *E. coli One Shot® TOP10* Chemically FIG. 8C).

- 5 The nucleotide sequence of the inserted DNA was determined. The pENTR/Xyn3 vector with the correct *xyn3* sequence was recombined with pTrex3g using a LR clonase® reaction protocol outlined by Invitrogen. The LR clonase reaction mixture was transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen), resulting in the final expression vector, pTrex3g/Xyn3 (FIG. 8D). The vector also contains the *Aspergillus nidulans*
10 *amdS* gene, encoding acetamidase, as a selectable marker for transformation of *T. reesei*. The expression cassette was amplified by PCR with primers SK745 and SK822 to generate product for transformation of *T. reesei*.

Forward Primer SK745: (5'-GAGTTGTGAAGTCGGTAATCC-3') (SEQ ID NO:129)

Reverse Primer SK822: (5'-CACGAAGAGCGGCGATTC-3') (SEQ ID NO:130)

15 C. Construction of the β -xylosidase Fv3A expression vector

[0377] The *F. verticillioides* β -xylosidase *fv3A* gene was amplified from a *F. verticillioides* genomic DNA sample using the primers MH124 and MH125.

Forward Primer MH124: (5'-CAC CCA TGC TGC TCA ATC TTC AG -3') (SEQ ID
NO:131)

- 20 Reverse Primer MH125: (5'-TTA CGC AGA CTT GGG GTC TTG AG -3') (SEQ ID
NO:132)

[0378] The PCR fragments were cloned into the Gateway ® Entry vector pENTR™/D-TOPO®, and transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen) resulting in the intermediate vector, pENTR-Fv3A (FIG. 8E). The nucleotide

- 25 sequence of the inserted DNA was determined. The pENTR-Fv3A vector with the correct *fv3A* sequence was recombined with pTrex6g (FIG. 8F) using a LR clonase® reaction protocol outlined by Invitrogen. The LR clonase reaction mixture was transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen), resulting in the final expression vector, pTrex6g/Fv3A (FIG. 8G) . The vector also contains a chlorimuron ethyl
30 resistant mutant of the native *T.reesei* acetolactate synthase (*als*) gene, designated *alsR*, which is used together with its native promoter and terminator as a selectable marker for transformation of *T. reesei* (WO2008/039370 A1). The expression cassette was PCR amplified with primers

SK1334, SK1335 and SK1299 to generate product for transformation of *T. reesei*.

Forward Primer SK1334: (5'-GCTTGAGTGTATCGTGTAAG -3') (SEQ ID NO:133)

Forward Primer SK1335: (5'-GCAACGGCAAAGCCCCACTTC -3') (SEQ ID NO:134)

Reverse Primer SK1299: (5'-GTAGCGGCCGCTCATCTCATCTCATCCATCC -3') (SEQ ID NO:135)

D. Construction of the β -xylosidase Fv43D expression cassette

[0379] For the construction of the *F. verticillioides* β -xylosidase Fv43D expression cassette, the *fv43D* gene product was amplified from a *F. verticillioides* genomic DNA sample using the primers SK1322 and SK1297. A region of the promoter of the endoglucanase gene *eglI* was amplified by PCR from a *T. reesei* genomic DNA sample extracted from strain RL-P37, using the primers SK1236 and SK1321. These two PCR amplified DNA fragments were subsequently fused together in a fusion PCR reaction using the primers SK1236 and SK1297. The resulting fusion PCR fragment was cloned into pCR-Blunt II-TOPO vector (Invitrogen) to give the plasmid TOPO Blunt/Peg11-Fv43D (FIG. 8H) and *E. coli One Shot® TOP10*

Chemically Competent cells (Invitrogen) were transformed using this plasmid. Plasmid DNA was extracted from several *E. coli* clones and confirmed by restriction digest.

Forward Primer SK1322: (5'-CACCATGCAGCTCAAGTTTCTGTC-3') (SEQ ID NO:136)

Reverse Primer SK1297: (5'-GGTTACTAGTCAACTGCCCGTTCTGTAGCGAG-3') (SEQ ID NO:137)

Forward Primer SK1236: (5'-CATGCGATCGCGACGTTTTGGTCAGGTCG-3') (SEQ ID NO:138)

Reverse Primer SK1321: (5'-GACAGAACTTGAGCTGCATGGTGTGGGACAACAAGAAGG-3') (SEQ ID NO:139)

[0380] The expression cassette was PCR amplified from TOPO Blunt/Peg11-Fv43D with primers SK1236 and SK1297 to generate product for transformation of *T. reesei*.

E. Construction of the α -arabinofuranosidase expression cassette

[0381] For the construction of the *F. verticillioides* α -arabinofuranosidase gene *fv51A* expression cassette, the *fv51A* gene product was amplified from *F. verticillioides* genomic DNA using the primers SK1159 and SK1289. A region of the promoter of the endoglucanase gene *eglI* was amplified by PCR from a *T. reesei* genomic DNA sample extracted from strain RL-P37, using the primers SK1236 and SK1262. These two PCR amplified DNA fragments were subsequently fused together in a fusion PCR reaction using the primers SK1236 and

SK1289. The resulting fusion PCR fragment was cloned into pCR-Blunt II-TOPO vector (Invitrogen) to give the plasmid TOPO Blunt/Peg11-Fv51A (FIG. 8I) and *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen) were transformed using this plasmid.

Forward Primer SK1159: (5'-CACCATGGTTCGCTTCAGTTCAATCCTAG-3') (SEQ ID NO:140)

Reverse Primer SK1289: (5'-GTGGCTAGAAGATATCCAACAC-3') (SEQ ID NO:141)

Forward Primer SK1236: (5'-CATGCGATCGCGACGTTTTGGTCAGGTCG-3') (SEQ ID NO:142)

Reverse Primer SK1262: (5'-GAACTGAAGCGAACCATGGTGTGGGACAACAAGAA GGAC-3') (SEQ ID NO:143)

[0382] The expression cassette was PCR amplified with primers SK1298 and SK1289 to generate product for transformation of *T. reesei*.

Forward Primer SK1298: (5'-GTAGTTATGCGCATGCTAGAC-3') (SEQ ID NO:144)

Reverse Primer SK1289: (5'-GTGGCTAGAAGATATCCAACAC-3') (SEQ ID NO:145)

F. Co-Transformation of T. reesei expression cassettes for β -glucosidase and endoxylanase

[0383] A *Trichoderma reesei* mutant strain, derived from RL-P37 (Sheir-Neiss, G *et al.* Appl. Microbiol. Biotechnol. 1984, 20:46-53), and selected for high cellulase production was co-transformed with the β -glucosidase expression cassette (*cbh1* promoter, *T. reesei* β -glucosidase1 gene, *cbh1* terminator, and *amdS* marker), and the endoxylanase expression cassette (*cbh1* promoter, *T. reesei xyn3*, and *cbh1* terminator) using PEG-mediated transformation (Penttila, M *et al.* Gene 1987, 61(2):155-64). Numerous transformants were isolated and examined for β -glucosidase and endoxylanase production. One transformant called *T. reesei* strain #229 was used for transformation with the other expression cassettes.

G. Co-transformation of T. reesei strain #229 with expression cassettes for two β -xylosidases and an α -arabinofuranosidase

[0384] *T. reesei* strain #229 was co-transformed with the β -xylosidase *fv3A* expression cassette (*cbh1* promoter, *fv3A* gene, *cbh1* terminator, and *alsR* marker), the β -xylosidase *fv43D* expression cassette (*egl1* promoter, *fv43D* gene, native *fv43D* terminator), and the *fv51A* α -arabinofuranosidase expression cassette (*egl1* promoter, *fv51A* gene, *fv51A* native terminator) using electroporation (see *e.g.* WO 08153712). Transformants were selected on Vogels agar plates containing chlorimuron ethyl (80 ppm). Vogels agar was prepared as follows, per liter.

50 x Vogels Stock Solution (recipe below) 20 mL

BBL Agar 20 g
 With deionized H₂O bring to 980 mL
 post-sterile addition:

5 50% Glucose 20 mL

50 x Vogels Stock Solution, per liter:

In 750 mL deionized H₂O, dissolve successively:

Na₃Citrate*2H₂O 125 g
 KH₂PO₄ (Anhydrous) 250 g
 10 NH₄NO₃ (Anhydrous) 100 g
 MgSO₄*7H₂O 10 g
 CaCl₂*2H₂O 5 g

Vogels Trace Element Solution (recipe below) 5 mL

d-Biotin 0.1 g
 15 With deionized H₂O, bring to 1 L

Vogels Trace Element Solution:

Citric Acid 50 g
 ZnSO₄*7H₂O 50 g
 Fe(NH₄)₂SO₄*6H₂O 10 g
 20 CuSO₄.5H₂O 2.5 g
 MnSO₄.4H₂O 0.5 g
 H₃BO₃ 0.5 g
 Na₂MoO₄.2H₂O 0.5 g

[0385] Numerous transformants were isolated and examined for β- xylosidase and L-α-
 25 arabinofuranosidase production. Transformants were also screened for biomass conversion
 performance according to the cob saccharification assay described in Example 1 (supra).
 Examples of *T. reesei* integrated expression strains described herein are H3A, 39A, A10A,
 11A, and G9A, which express all of the genes for *T. reesei* beta-glucosidase 1, *T. reesei*
 Xyn3, Fv3A, Fv51A, and Fv43D, at different ratios. Other integrated *T. reesei* strains
 30 include those wherein most of the genes for *T. reesei* beta-glucosidase 1, *T. reesei* Xyn3,
 Fv3A, Fv51A, and Fv43D, were expressed at different ratios. For example, one lacked
 overexpressed *T. reesei* Xyn3; another lacked Fv51A, as determined by Western Blot; two
 others lacked Fv3A, one lacked overexpressed Bgl1 (e.g. strain H3A-5).

H. Composition of *T. reesei* integrated strain H3A

35 [0386] Fermentation of the *T. reesei* integrated strain H3A yields the following proteins *T.*
reesei Xyn3, *T. reesei* Bgl 1, Fv3A, Fv51A, and Fv43D, at ratios determined as described
 herein and shown in FIG. 9.

I. Protein Analysis by HPLC

[0387] Liquid chromatography (LC) and mass spectroscopy (MS) were performed to
 40 separate, identify, and quantify the enzymes contained in fermentation broths. Enzyme

samples were first treated with a recombinantly expressed endoH glycosidase from *S. plicatus* (e.g., NEB P0702L). EndoH was used at a ratio of 0.01-0.03 µg endoH protein per µg sample total protein and incubated for 3 h at 37°C, pH 4.5-6.0 to enzymatically remove N-linked glycosylation prior to HPLC analysis. Approximately 50 µg of protein was then
5 injected for hydrophobic interaction chromatography using an Agilent 1100 HPLC system with an HIC-phenyl column and a high-to-low salt gradient over 35 min. The gradient was achieved using high salt buffer A: 4 M ammonium sulphate containing 20 mM potassium phosphate pH 6.75 and low salt buffer B: 20 mM potassium phosphate pH 6.75. Peaks were detected with UV light at 222 nm and fractions were collected and identified by mass
10 spectroscopy. Protein concentrations are reported as the percent of each peak area relative to the total integrated area of the sample.

J. Effect of addition of purified proteins to the fermentation broth of T. reesei integrated strain H3A on saccharification of dilute ammonia pretreated corn cob

[0388] Purified proteins (and one unpurified protein) were serially diluted from stock

15 solution and added to a fermentation broth of *T. reesei* integrated strain H3A to determine their benefit to saccharification of pretreated biomass. Dilute ammonia pretreated corn cob was loaded into microtiter plate (MTP) wells at 20% solids (w/w) (~5 mg of cellulose per well), pH 5. H3A protein (in the form of fermentation broth) was added to each well at 20 mg protein/g cellulose. Volumes of 10, 5, 2, and 1 µL of each of the diluted proteins (FIG. 10)
20 were added into individual wells, and water was added such that the liquid addition to each well was a total of 10 µL. Reference wells included additions of either 10 µL water or dilutions of additional H3A fermentation broth. The MTP were sealed with foil and incubated at 50°C with 200 RPM shaking in an Innova incubator shaker for three days. The samples were quenched with 100 µL of 100 mM glycine pH 10. The quenched samples were covered
25 with a plastic seal and centrifuged 3000 RPM for 5 min at 4°C. An aliquot (5 µL) of the quenched reactions was diluted with 100 µL of water and the concentration of glucose produced in the reactions was determined using HPLC. The glucose data was plotted as a function of the protein concentration added to the 20 mg/g of H3A (the concentrations of the protein additions were variable due to different starting concentrations and additions by
30 volume). Results are shown in FIGs. 11A-11D.

Example 3: Construction of T. reesei strains

A. Construction of and screening for T. reesei strain H3A/EG4#27

[0389] An expression cassette containing the *T. reesei egII* (also termed “Cel 7B”)

promoter, *T. reesei* *eg4* (also termed “TrEG4”, or “Cel 61A”) open reading frame, and *cbh1* (Cel 7A) terminator sequence (FIG. 12A) from *Trichoderma reesei*, and *sucA* selectable marker (see, Boddy et al., Curr. Genet. 1993, 24:60-66) from *Aspergillus niger* was cloned into pCR Blunt II TOPO (Invitrogen) (FIG. 12B).

- 5 [0390] The expression cassette *Pegl1-eg4-sucA* was amplified by PCR with the primers: SK1298: 5'-GTAGTTATGCGCATGCTAGAC-3' (SEQ ID NO:146)
214: 5'-CCGGCTCAGTATCAACCACTAAGCACAT-3' (SEQ ID NO:147)

[0391] *Pfu Ultra II* (Stratagene) was used as the polymerase for the PCR reaction. The products of the PCR reaction were purified with the QIAquick PCR purification kit (Qiagen)
10 as per the manufacturer's protocol. The products of the PCR reaction were then concentrated using a speed vac to 1-3 µg/µL. The *T. reesei* host strain to be transformed (H3A) was grown to full sporulation on potato dextrose agar plates for 5 d at 28°C. Spores from 2 plates were harvested with MilliQ water and filtered through a 40 µM cell strainer (BD Falcon). Spores were transferred to a 50 mL conical tube and washed 3 times by repeated centrifugation with
15 50 mL water. A final wash with 1.1 M sorbitol solution was carried out. The spores were resuspended in a small volume (less than 2 times the pellet volume) using 1.1 M sorbitol solution. The spore suspension was then kept on ice. Spore suspension (60 µL) was mixed with 10-20 µg of DNA, and transferred into the electroporation cuvette (E-shot, 0.1 cm standard electroporation cuvette from Invitrogen). The spores were electroporated using the
20 Biorad Gene Pulser Xcell with settings of 16 kV/cm, 25 µF, 400 Ω. After electroporation, 1 mL of 1.1.M sorbitol solution was added to the spore suspension. The spore suspension was plated on Vogel's agar (see example 2G), containing 2% sucrose as the carbon source.

[0392] The transformation plates were incubated at 30°C for 5-7 d. The initial transformants were restreaked onto secondary Vogel's agar plates with sucrose and grown at
25 30°C for an additional 5-7 d. Single colonies growing on secondary selection plates were then grown in wells of microtiter plates using the method described in WO/2009/114380. The supernatants were analyzed on SDS-PAGE to check for expression levels prior to saccharification performance screening.

[0393] A total of 94 transformants overexpressed EG4 in strain H3A. Two H3A control
30 strains were grown in microtiter plates along with the H3A/EG4 strains. Performance screening of *T. reesei* strains expressing EG4 protein was performed using ammonia pretreated corncob. The dilute ammonia pretreated corncob was suspended in water and

adjusted to pH 5.0 with sulfuric acid to achieve 7% cellulose. The slurry was dispensed into a flat bottom 96 well microtiter plate (Nunc, 269787) and centrifuged at 3,000 rpm for 5 min.

[0394] Corncob saccharification reactions were initiated by adding 20 μ L of H3A or H3A/EG4 strain culture broth per well of substrate. The corncob saccharification reactions were sealed with aluminum (E&K scientific) and mixed for 5 min at 650 rpm, 24°C. The plate was then placed in an Innova incubator at 50°C and 200 rpm for 72 h. At the end of 72-h saccharification, the reactions were quenched by adding 100 μ L of 100 mM glycine, pH 10.0. The plate was then mixed thoroughly and centrifuged at 3,000 rpm for 5 min. Supernatant (10 μ L) was added to 200 μ L of water in an HPLC 96-well microtiter plate (Agilent, 5042-1385).

Glucose, xylose, cellobiose and xylobiose concentrations were measured by HPLC using an Aminex HPX-87P column (300 mm x 7.8 mm, 125-0098) pre-fitted with guard column.

[0395] The screening on corncob identified the following H3A/EG4 strains as having improved glucan and xylan conversion compared to the H3A control strains: 1, 2, 3, 4, 5, 6, 14, 22, 27, 43, and 49 (FIG. 13).

[0396] Select H3A/EG4 strains were re-grown in shake flasks. A total of 30 mL of protein culture filtrate was collected per shake flask per strain. The culture filtrates were concentrated 10-fold using 10 kDa membrane centrifugal concentrators (Sartorius, VS2001) and the total protein concentration was determined by BCA as described in Example 1C. A corncob saccharification reaction was performed using 2.5, 5, 10, or 20 mg protein from H3A/EG4 strain samples per g of cellulose per well of corncob substrate. An H3A strain produced at 14 L fermentation scale and a previously identified low performance sample (H3A/EG4 strain #20) produced at shake flask scale were included as controls. The saccharification reactions were carried out as described in Example 4 (below). Increased glucan conversion with increased protein dose was observed with culture supernatant from all of the EG4 expressing strains (FIG. 14). *T. reesei* integrated strain H3A/EG4#27 was used in additional saccharification reactions, and the strain was purified by streaking a single colony onto a potato dextrose plate from which a single colony was isolated.

Example 4: Range of *T. reesei* EG4 concentrations for improved saccharification of dilute ammonia pretreated corncob

[0397] To determine preferred dosing, hydrolysis of dilute ammonia pretreated corncob (25% solids, 8.7% cellulose, 7.3% xylan) was conducted at pH 5.3 using fermentation broth from either *T. reesei* integrated strain H3A/EG4 #27 or H3A with purified EG4 added to the reaction mix. The total loading of *T. reesei* integrated strain H3A/EG4 #27 or H3A was 14

mg protein per gram of glucan (G) and xylan (X).

[0398] The reaction mix (total mass 5 g) was loaded into 20 mL scintillation vials in a total reaction volume of 5 mL according to the dosing chart in FIGs. 15, 17A and 17B.

[0399] The set up for experiment 1 is shown in FIG. 15. MilliQ Water and 6 N Sulfuric acid were mixed in a conical tube and added to the respective vials and the vials were swirled to mix the contents. Enzymes samples were added to the vials and the vials incubated for 6 d at 50°C. At various time points, 100µL of sample was removed from the vialss diluted with 900 µL 5mM sulfuric acid, vortexed, centrifuged and the supernatant was used to measure the concentrations of soluble sugars using HPLC. The results of glucan and xylan conversion are shown in FIGs. 16A and 16B, respectively.

[0400] The set up for experiment 2 is shown in FIG. 17A. To further determine the preferred EG4 concentration, saccharification of dilute ammonia corncob (25% solids, 8.7% cellulose, 7.3% xylan) was conducted at pH 5.3 using fermentation broth from either *T. reesei* integrated strain H3A/EG4 #27 or H3A with purified EG4 added (ranging from 0.05 to 1.0 mg protein/g G+X) to the reaction mix. The total loading of *T. reesei* integrated strain H3A/EG4 #27 or H3A was 14 mg protein/g glucan + xylan. The experimental results are shown in FIG. 18A.

[0401] The set up for experiment 3 is shown in FIG. 17B. To pinpoint the preferred concentration range of *T. reesei* Eg4 yet further, dilute ammonia corncob (25% solids, 8.7% cellulose, and 7.3% xylan) was hydrolyzed at pH 5.3 using *T. reesei* integrated strain H3A/EG4 #27 or H3A with purified EG4 added at concentrations ranging from 0.1-0.5 mg protein/g G+X. The total loading of *T. reesei* integrated strain H3A/EG4 #27 or H3A was 14 mg protein per g of glucan and xylan.

[0402] Results are shown in FIG. 18B.

Example 5: Effect of *T. reesei* Eg4 on saccharification of dilute ammonia pretreated corn stover at different solid loadings

[0403] Dilute ammonia pre-treated corn stover was incubated with fermentation broth from *T. reesei* integrated strain H3A or H3A/EG4#27 (14 mg protein/g glucan and xylan) at 7, 10, 15, 20 and 25% solids (%S) for three days at 50°C, pH 5.3 (5 g total wet biomass in 20 mL vials). The reactions were carried out as described in Example 4 above. Glucose and xylose were analyzed by HPLC. Results are shown in FIG. 19. All samples up to 20% solids were visibly liquefied on day 1.

Example 6: Effect of overexpression of *T. reesei* EG4 on hydrolysis of dilute ammonia pretreated corncob

[0404] The effect of overexpression of *T. reesei* Eg4 in strain H3A on saccharification of dilute ammonia pretreated corncob was tested using fermentation broths from strains H3A/EG4 # 27 and H3A. Corncob saccharification at 3 g scale was performed in 20 mL glass vials as follows. Enzyme preparation, 1 N sulfuric acid and 50 mM pH 5.0 sodium acetate buffer (with 0.01% sodium azide and 5 mM MnCl₂) were added to give a final slurry of 3 g total reaction, 22% dry solids, pH 5.0 with enzyme loadings varying between 1.7 and 21.0 mg total protein per gram Glucan + Xylan. All saccharification vials were incubated at 48°C with 180 rpm rotation. After 72 h, 12 mL of filtered MilliQ water was added to each vial to dilute the entire saccharification reaction 5-fold. The samples were centrifuged at 14,000 x g for 5 min, then filtered through a 0.22 µm nylon filter (Spin-X centrifuge tube filter, Corning Incorporated, Corning, NY) and further diluted 4-fold with filtered MilliQ water to create a final 20X dilution. 20 µL injections were analyzed by HPLC to measure the sugars released.

[0405] Overexpression or addition of *T. reesei* Eg4 led to enhanced xylose and glucose monomer release as compared to H3A alone (FIGs. 20 and 21). Addition of H3A/EG4#27 at different doses led to an increased yield of xylose as compared to strain H3A, or compared to Eg4 + a constant 1.12 mg Xyn3 per g Glucan + Xylan (FIG. 20).

[0406] Addition of H3A/EG4#27 at different doses led to an increased yield of glucose compared to strain H3A or compared to Eg4 + a constant 1.12 mg Xyn3 per g Glucan + Xylan (FIG. 21).

[0407] The effect of *T. reesei* Eg4 on total fermentable monomer (xylose, glucose and arabinose) release by integrated strains H3A/EG4# 27 or H3A is illustrated in the FIG. 22. The H3A/EG4#27 integrated strain led to enhanced total fermentable monomer release compared to the integrated strain H3A, or compared to Eg4 + 1.12 mg Xyn3/g Glucan + Xylan.

Example 7: Purified *T. reesei* EG4 leads to glucose release in dilute ammonia pretreated corncob

[0408] The effect of purified *T. reesei* Eg4 on the concentration of sugars released was tested using 1.05 g dilute ammonia pretreated corncob in the presence or absence of 0.53 mg Xyn3 per g Glucan + Xylan. The experiments were performed as described in Example 6. Results are shown in FIG. 23. The data indicate that purified *T. reesei* Eg4 leads to release of

glucose monomer without the action of other cellulases such as endoglucanases, cellobiohydrolases and β -glucosidases.

[0409] Saccharification experiments were also conducted using dilute ammonia pretreated corncob with purified Eg4 added alone (no Xyn3 added). 3.3 μ L of purified Eg4 (15.3 mg/mL) was added to 872 μ L 50 mM, pH 5.0 sodium acetate buffer (included 0.01% sodium azide and 5 mM MnCl_2), 165 mg of dilute ammonia pretreated corncob (67.3% dry solids, 111 mg dry solids added) and 16.5 μ L of 1 N sulfuric acid in 5 mL vials. The vials were incubated at 48°C and rotated at 180 rpm. Periodically, 20 μ L aliquots were removed, diluted 10-fold with filter sterilized double distilled water and filtered through a nylon filter before analysis for glucose released on a Dionex Ion Chromatography system. Authentic glucose solutions were used as external standards. Results are shown in FIG. 24, indicating that addition of purified Eg4 leads to release of glucose monomer from dilute ammonia pretreated corncobs over 72 h incubation at 48°C in the absence of other cellulases or endoxylanase.

Example 8: Saccharification performance of *T. reesei* integrated strains H3A and H3A/EG4 #27 on various substrates

[0410] In this experiment, fermentation broth from *T. reesei* integrated strain H3A or H3A/EG4#27, dosed at 14 mg protein per g of glucan + xylan, was tested for saccharification performance on different substrates including: dilute ammonia pretreated corncob, washed dilute ammonia pretreated corncob, ammonia fiber expanded corn stover (AFEX CS), Steam Expanded Sugarcane Bagasse (SEB), and Kraft-pretreated paper pulps FPP27 (Softwood Industrial Unbleached Pulp delignified-Kappa 13.5, Glucan 81.9%, Xylan 8.0%, Klason Lignin 1.9%), FPP-31 (Hardwood Unbleached Pulp delignified-Kappa 10.1, Glucan 75.1%, Xylan 19.1%, Klason Lignin 2.2%), and FPP-37 (Softwood Unbleached Pulp air dried-Kappa 82, Glucan 71.4%, Xylan 8.7%, Klason Lignin 11.3%).

[0411] The saccharification reactions were set up in 25 mL glass vials with final mass of 10 g in 0.1 M Sodium Citrate Buffer, pH 5.0 and incubated at 50°C, 200 rpm for 6 d. At the end of 6 d, 100 μ L aliquots were diluted 1:10 in 5 mM sulfuric acid and the samples analyzed by HPLC to determine glucose and xylose formation. Results are shown in FIG. 25.

Example 9: Effect of *T. reesei* EG4 on saccharification of acid pretreated corn stover

[0412] The effect of Eg4 on saccharification of acid pretreated corn stover was tested. Corn stover pretreated with dilute sulfuric acid (Schell, DJ, et al., *Appl. Biochem. Biotechnol.* 2003, 105(1-3):69-85) was obtained from NREL, adjusted to 20% solids and conditioned to a pH 5.0 with the addition of soda ash solution. Saccharification of the pretreated substrate was

performed in a microtiter plate using 20% total solids. Total protein in the fermentation broths was measured by the Biuret assay (see Example 1 above). Increasing amounts of fermentation broth from *T. reesei* integrated strains H3A/EG4 #27 and H3A were added to the substrate and saccharification performance was measured following incubation at 50°C, 5 d, 200 RPM shaking. Glucose formation (mg/g) was measured using HPLC. Results are shown in FIG. 26.

Example 10: Saccharification performance of *T. reesei* integrated strains H3A and H3A/EG4#27 on dilute ammonia pretreated corn leaves, stalks, and cobs

[0413] Saccharification performance of *T. reesei* integrated strains H3A and H3A/EG4#27 was compared on dilute ammonia pretreated corn stover leaves, stalks, or cobs. Pretreatment was performed as described in WO06110901A. Five (5) g total mass (7% solids) was hydrolyzed in 20 mL vials at pH 5.3 (pH adjusted with 6 N H₂SO₄) using 14 mg protein per g of glucan+xylan. Saccharification reactions were carried out at 50°C and samples analyzed by HPLC for glucose and xylose released on day 4. Results are shown in FIG. 27.

Example 11: Saccharification performance on dilute ammonia pretreated corncob in response to overexpressed EG4 from *T. reesei*

[0414] Saccharification reactions at 3 g scale were performed using dilute ammonia pretreated corncob. Sufficient pretreated cob preparation was measured into 20 mL glass vials to give 0.75 g dry solid. Enzyme preparation, 1 N sulfuric acid and 50 mM pH 5.0 sodium acetate buffer (with 0.01% sodium azide) were added to give final slurry of 3 g total reaction, 25% dry solids, pH 5.0. Extra cellular protein (fermentation broth) from the *T. reesei* integrated strain H3A was added at 14 mg protein/ g (glucan+xylan) either with or without an additional 5% of the 14 mg protein load as the unpurified culture supernatant from a *T. reesei* strain ($\Delta cbh1 \Delta cbh2 \Delta eg1 \Delta eg2$) (See International publication WO 05/001036) over expressing Eg4. The saccharification reactions were incubated for 72 h at 50°C. Following incubation, the reaction contents were diluted 3-fold, filtered and analyzed by HPLC for glucose and xylose concentration. The results are shown in FIG. 28. Addition of Eg4 protein in the form of extracellular protein from a *T. reesei* strain over expressing Eg4 to H3A substantially increased the release of monomer glucose and slightly increased the release of monomer xylose.

Example 12: Saccharification performance of strain H3A/EG4#27 on ammonia pretreated switchgrass

[0415] The saccharification performance of strain H3A/EG4#27 on ammonia pretreated

switchgrass (International Patent Publication WO06110901A) at increasing protein doses was compared to that of strain H3A (18.5% solids). Pretreated switchgrass preparations were measured into 20 mL glass vials to give 0.925 g of dry solid. 1 N sulfuric acid and 50 mM pH 5.3 sodium acetate buffer (with 0.01% sodium azide) were added to give final slurry of 5 grams total reaction. The enzyme dosages of H3A tested were 14, 20, and 30 mg/g (glucan + xylan); and the dosages of H3A-EG4 #27 were 5, 8, 11, 14, 20, and 30 mg/g (glucan + xylan). The reactions were incubated at 50°C for 3 d. Following incubation, the reaction contents were diluted 3-fold, filtered and analyzed by HPLC for glucose and xylose concentration. The conversion of glucan and xylan were calculated based on the composition of the switchgrass substrate. The results (FIG. 29) indicate that the performance of H3A-EG4 #27 is more effective for glucan conversion than H3A at the same enzyme dosages.

Example 13: Effect of *T. reesei* EG4 additions on corncob saccharification and on CMC and cellobiose hydrolysis

A. *Corncob* saccharification:

[0416] Dilute ammonia pretreated corncob was adjusted to 20% solids, 7% cellulose and 65 mg was dispensed per well in a microtiter plate. Saccharification reactions were initiated by adding 35 μ L of 50 mM sodium acetate (pH 5.0) buffer containing *T. reesei* CBH1 at 5 mg protein/g glucan (final) and the relevant enzymes (CBH1 or Eg4), at final concentrations of 0, 1, 2, 3, 4 and 5 mg/g glucan. An Eg4 control received only EG4 at the same doses and as such, the total added protein in these wells was less. The microtiter plates were sealed with an aluminum plate seal (E&K scientific) and mixed for 2 min at 600 rpm, 24°C. The plate was then placed in an Innova incubator at 50°C and 200 rpm for 72 h.

[0417] At the end of 72-h saccharification, the plate was quenched by adding 100 μ L of 100 mM glycine, pH 10.0. The plate was then centrifuged at 3000 rpm for 5 min. Supernatant (20 μ L) was added to 100 μ L of water in HPLC 96 well microtiter plate (Agilent 5042-1385). Glucose and cellobiose concentrations were measured by HPLC using Aminex HPX-87P column (300 mm x 7.8 mm, 125-0098) pre-fitted with guard column. %glucan conversion was calculated by $100 \times (\text{mg cellobiose} + \text{mg glucose}) / \text{total glucan in substrate}$ (FIG. 30).

B. CMC hydrolysis:

[0418] Carboxymethylcellulose (CMC, Sigma C4888) was diluted to 1% with 50 mM Sodium Acetate, pH 5.0. Hydrolysis reactions were initiated by separately adding each of three *T. reesei* purified enzymes – EG4, EG1 and CBH1 at final concentrations of 20, 10, 5, 2.5, 1.25 and 0 mg/g to 100 μ L of 1% CMC in a 96-well microtiter plate (NUNC #269787).

Sodium acetate, pH 5.0 50 mM was added to each well to a final volume of 150 μ L. The CMC hydrolysis reactions were sealed with an aluminum plate seal (E&K scientific) and mixed for 2 min at 600 rpm, 24°C. The plate was then placed in an Innova incubator at 50°C and 200 rpm for 30 min.

- 5 [0419] At the end of 30 min. incubation, the plate was put in ice water for 10 min. to stop the reaction, and samples were transferred to eppendorf tubes. To each tube was added 375 μ L of dinitrosalicylic acid (DNS) solution (see below). Samples were then boiled for 10 min and O.D was measured at 540 nm by SpectraMAX 250 (Molecular Devices). Results are shown in FIG. 31.

10 DNS SOLUTION:

40 g 3,5-Dinitrosalicylic acid (Sigma, D0550)

8 g Phenol

2 g Sodium sulfite (Na_2SO_3)

800 g Na-K tartarate (Rochelle salt)

- 15 Add all the above to 2 L of 2% NaOH

Stir overnight, covered with aluminum foil

Add distilled deionized water to a final volume of 4 L

Mix well

Store in a dark bottle, refrigerated

- 20 C. Cellobiose hydrolysis

[0420] Cellobiose was diluted to 5 g/L with 50 mM Sodium Acetate, pH 5.0. Hydrolysis reactions were initiated by separately adding each of two enzymes – EG4 and BGL1 at final concentrations of 20, 10, 5, 2.5, and 0 mg/g to 100 μ L cellobiose solution at 5 g/L. Sodium acetate, pH 5.0 was added to each well to a final volume of 120 μ L. The reaction plates were sealed with an aluminum plate seal (E&K scientific) and mixed for 2 min at 600 rpm, 24°C. The plate was then placed in an Innova incubator at 50°C and 200 rpm for 2 h.

- [0421] At the end of the 2 h hydrolysis step, the plate was quenched by adding 100 μ L of 100 mM glycine, pH 10.0. The plate was then centrifuged at 3000 rpm for 5 min. Glucose concentration was measured by ABTS (2,2'-azino-bis 3-ethylbenzothiazoline-6-sulfonic acid) assay (Example 1). Ten (10) μ L of supernatant was added to 90 μ L ABTS solution in a 96-well microtiter plate (Corning costar 9017 EIA/RIA plate, 96 well flat bottom, medium binding). OD 420 nm was measured by SpectraMAX 250, Molecular Devices. Results are shown in FIG. 32.

35 Example 14: Purified EG4 improves glucose production from dilute ammonia pretreated corn cob when mixed with various cellulase mixtures

[0422] The effect of purified Eg4 combined with purified cellulases (*T. reesei* EG1, EG2,

CBH1, CBH2, and Bgl1) on the concentration of sugars released was tested using 1.05 g dilute ammonia pretreated corncob in the presence of 0.53 mg *T. reesei* Xyn3 per g of Glucan + Xylan. 1.06-g reactions were set up in 5 mL vials containing 0.111 g dry cob solids (10.5% solids). Enzyme preparation (FIG. 33), 1 N sulfuric acid and 50 mM pH 5.0 sodium acetate buffer (with 0.01% sodium azide and 5 mM MnCl₂) were added to give the final reaction weight. The reaction vials were incubated at 48°C with 180 rpm rotation. After 72 h, filtered MilliQ water was added to dilute each saccharification reaction by 5-fold. The samples were centrifuged at 14,000xg for 5 min, then filtered through a 0.22 µm nylon filter (Spin-X centrifuge tube filter, Corning Incorporated, Corning, NY) and further diluted 4-fold with filtered Milli-Q water to create a final 20X dilution. Twenty (20) µL injections were analyzed by HPLC to measure the sugars released (glucose, cellobiose, and xylose).

[0423] FIG. 34 shows glucose (A), glucose + cellobiose (B), or xylose (C) produced with each combination. Purified Eg4 improved the performance of individual cellulases and mixtures. When all of the purified cellulases were present, addition of 0.53 mg Eg4 per g Glucan + Xylan improved the conversion by almost 40%. Improvement was also seen when Eg4 was added to a combination of CBH1, Egl1 and Bgl1. When individual cellulases were present with the cob, the absolute amounts of total glucose release were substantially lower than resulted from the experiment wherein combinations of cellulases were present with the cob, but in each case, the percent improvement in the presence of Eg4 was significant.

Addition of *T. reesei* Eg4 to purified cellulases resulted in the following percent improvements in total Glucose release-Bgl1 (121%), Egl2 (112%), CBH2 (239%) and CBH1 (71%). This shows that Eg4 had a significant and broad effect to improve cellulase performance on biomass.

Example 15: Effects Observed When EG4 was Mixed with CBH1, CBH2, and EG2 – Substrate: Dilute Ammonia Pretreated Corncob

[0424] Dilute ammonia pretreated corncob saccharification reactions were prepared by adding enzyme mixtures as follows to corncob (65 mg per well of 20% solids, 7% cellulose) in 96-well MTPs (VWR). Eighty (80) µL of 50 mM sodium acetate (pH 5.0), 1 mg Bgl1/g glucan, and 0.5 mg Xyn3/g glucan background were also added to all wells.

[0425] To test the effect of mixing Eg4 individually with CBH1, CBH2 and EG2, each of CBH1, CBH2, and EG2 was added at 0, 1.25, 2.5, 5, 10 and 20 mg/g glucan, and EG4 was added at concentrations of 20, 18.75, 17.5, 15, 10 and 0 mg/g glucan to the respective wells, making the total proteins in individual wells 20 mg/g glucan. The control wells received only

CBH1 or CBH2 or EG2 or EG4 at the same doses, as such the total added proteins in these wells were less than 20 mg/g.

[0426] To test the effect of Eg4 on combinations of cellulases, mixtures of CBH1, CBH2 and EG2 at different ratios (*see*, FIG. 35) were added at 0, 1.25, 2.5, 5, 10 and 20 mg protein/g glucan, and EG4 was added to the mixtures at concentrations of 20, 18.75, 17.5, 15, 10 and 0 mg protein/g glucan, such that the total proteins in individual wells was 20 mg protein/g glucan. As above, control wells received only one added protein so the total protein addition was less than 20 mg protein/g.

[0427] The corncob saccharification reactions were sealed with an aluminum plate seal (E&K scientific) and mixed for 2 min at 600 rpm, 24°C. The plate was then placed in an Innova 44 incubator shaker (New Brunswick Scientific) at 50 °C and 200 rpm for 72 h. At the end of the 72-h saccharification step, the plate was quenched by adding 100 µL of 100 mM glycine, pH 10.0. The plate was then centrifuged at 3000 rpm for 5 min (Rotanta 460R Centrifuge, Hettich Zentrifugen). Twenty (20) µL of supernatant was added to 100 µL of water in an HPLC 96-well microtiter plate (Agilent, 5042-1385). Glucose and cellobiose concentrations were measured by HPLC using an Aminex HPX-87P column (300 mm x 7.8 mm, 125-0098) and guard column (BioRad).

[0428] The results were indicated in the table of FIG. 36, wherein the glucan conversion (%) is defined as $100 \times (\text{glucose} + \text{cellulobiose}) / \text{total glucan}$.

[0429] This experiment indicates that Eg4, when added to a CBH1, CBH2 and/or EG2, was beneficial in improving saccharification of dilute ammonia pretreated corncob. Moreover, the highest improvement was observed when Eg4 and the other enzyme (CBH1, CBH2, or EG2) were added to the saccharification mixture in an equal amount. It was also observed that the effect of Eg4 is substantial on the CBH1 and CBH2 mixture. The optimum improvement by Eg4 was observed when the amount of Eg4 to CBH1 and CBH2 was 1:1.

Example 16: EG4 Improves Saccharification Performance of Various Cellulase Compositions

[0430] The total protein concentration of commercial cellulase enzyme preparations Spezyme® CP, Accellerase®1500, and Accellerase®DUET (Genencor Division, Danisco US) were determined by the modified Biuret assay (described herein).

[0431] Purified *T. reesei* EG4 was added to each enzyme preparation, and the samples were then assayed for saccharification performance using a 25% solids loading of ammonia pretreated corncob, at a dose of 14 mg of total protein per g of substrate glucan and xylan (5

mg EG4 per g of glucan and xylan, plus 9 mg whole cellulase per g of glucan and xylan).

The saccharification reaction was carried out using 5 g of total reaction mixture in a 20 mL vial at pH 5, with incubation at 50°C in a rotary shaker set to 200 rpm for 7 d. The saccharification samples were diluted 10x with 5 mM sulfuric acid, filtered through a 0.2 µm filter before injection into the HPLC. HPLC analysis was performed using a BioRad Aminex HPX-87H ion exclusion column (300 mmx7.8 mm).

[0432] Substitution of purified EG4 into whole cellulases improved glucan conversion in all tested cellulase products as illustrated in FIG. 40. As illustrated in FIG. 41, xylan conversion did not appear to be affected by the Eg4 substitution.

10 **Example 17: Reduction of Viscosity in Biomass Saccharification**

[0433] Biomass used in this experiment was Inbicon acidified steam-expansion pretreated wheat straw, with the following composition (Table 2):

Component ID	Inbicon wheat straw
	Mean
Glucan	55.0%
Xylan	5.0%
Galactan	
Arabinan	
Mannan	
Klason Lignin	
Acid soluble lignin	31.0%
Ash	4.0%
Starch	
Mass Balance Closure	95.0%

[0434] The pre-treated wheat straw was diluted into water and pH-adjusted with sulfuric

acid to pH5.0, and a solid level of 10.5% of that was mixed with, in a first sample, a fermentation broth of a *T. reesei* H3A strain (FIG. 9) at a total protein concentration of 20.5 mg protein/g cellulose in the biomass substrate at 50°C, or in a second sample, the fermentation broth of *T. reesei* H3A (FIG. 9) at a total protein concentration of 18.5 mg protein/g cellulose in the biomass substrate, and 2 mg/g cellulose of purified *T. reesei* Eg4.

Viscosity reduction was measured using a Brookfield viscometer (Brookfield Engineering, Inc), monitoring viscosity change up to about 6 h. Results are indicated in FIG. 42.

20 **Example 18: Reduction of Viscosity in Biomass Saccharification**

[0435] Biomass used in this experiment was dilute acid pretreated corn stover from NREL (unwashed PCS).

[0436] The unwashed pretreated corn stover was mixed, at a temperature of 50°C, pH of

5.0, and a solid level of 20% dry solids with, in a first sample, a fermentation broth of a *T. reesei* H3A strain (FIG. 9) at a total protein concentration of 20 mg/g cellulose in the biomass substrate, and in a second sample, a fermentation broth of *T. reesei* H3A/Eg4 #27 integrated strain, also at 20 mg/g cellulose. Viscosity reduction was measured using a Brookfield viscometer (Brookfield Engineering, Inc.), monitoring viscosity change for up to over 160 h. The results are indicated in FIG. 43.

Example 19: Reduction of Viscosity in Biomass Saccharification

[0437] Biomass used in this experiment was dilute ammonia pretreated corncob.

[0438] The dilute ammonia pretreated corncob was mixed with enzyme compositions at two solid loading conditions: 25% dry solids and 30% dry solids. Specifically, the pretreated biomass was mixed at 50°C and pH 5.0 with 14 mg protein/g cellulose from a fermentation broth of either a *T. reesei* H3A (FIG. 9) or H3A/Eg4 #27 strain. Viscosity reduction was measured using a Brookfield Viscometer (Brookfield Engineering, Inc.). The results are indicated in FIG. 44.

Example 20: Determining the effects of various cellulases on viscosity reduction and glucose production in saccharification process

[0439] This study used various viscosity reducing enzymes, such as OPTIMASH™ BG, OPTIMASH™ TBG, OPTIMASH™ VR; or beta-glucosidase such as Accellerase® BG, in the presence of Accellerase® DUET in the saccharification process and determined the effects of these viscosity reducing enzymes in glucose production and viscosity reduction. Enzyme composition produced from H3A/EG4 integrated strain #27 was also included. Accellerase® 1500, Accellerase® DUET, Accellerase® BG, OPTIMASH™ BG, OPTIMASH™ TBG, and OPTIMASH™ VR were products available from Danisco US Inc., Genencor.

[0440] Pretreated wheat straw as described above was used. The composition analysis was performed and is listed in Table 2 (see Example 17).

[0441] The saccharification process was performed by incubating the pretreated wheat straw (25% dry matter) with various enzymes in reaction chambers. See, Larsen et al., The IBUS Process- Lignocellulosic Bioethanol Close to A commercial Reality, (2008) Chem. Eng. Tech. 31(5):765-772. The experimental conditions are shown in Tables 3 and 4. In each chamber, the total mass was 10 kg. The initial pH of the wheat straw was about 3.50 and was adjusted by adding Na₂CO₃ to pH 5.0. Glucose concentration was measured over time and cellulose conversion was calculated.

Table 3.

Experimental condition	Enzymes	Cellulase Loading	Viscosity Enzyme
		mL/g cellulose	g/kg dry matter
1	Accellerase® 1500batch 1	0.22	0
2	Accellerase® DUET	0.15	0
3	Accellerase® DUET	0.25	0
4	Accellerase® DUET + Optimash™ BG	0.15	6
5	Accellerase® DUET + Optimash™ TBG	0.15	6
6	Accellerase® DUET + Optimash™ VR	0.15	6

Table 4.

Experimental condition	Enzymes	Cellulase Loading	Viscosity Enzyme
		mL/g cellulose	g/kg dry matter
7	Accellerase® 1500 (batch 1)	0.22	0
8	Accellerase® 1500 (batch 2)	0.22	0
9	Accellerase® DUET	0.15	0
10	Accellerase® DUET + Accellerase® BG	0.15	0.1
11	Accellerase® DUET + Accellerase® BG	0.15	6
12	H3A/Eg4#27	0.15	0

[0442] Experimental conditions 1-6 were conducted on the first day ("Day 1"), and experimental conditions 7-12 were conducted on the second day ("Day 2").

- 5 **[0443]** The glucose concentration was measured after 6 hour saccharification for each experimental condition. Accellerase® DUET at 0.25 mL/g cellulose resulted in 40.8 g glucose/kg after 6-h saccharification. See FIG. 45. The glucose concentration for Accellerase® DUET + OPTIMASH BG (or TBG) (0.15 + 6) (*i.e.*, 0.15 mL Accellerase® DUET/g cellulose + 6 g OPTIMASH BG (or TBG) / kg dry matter) was similar to the
- 10 glucose concentration for Accellerase® 1500 at 0.22 mL/g cellulose. See FIG. 45. The glucose concentration for Accellerase® DUET + Accellerase BG at 0.15 + 6 (*i.e.*, 0.15 mL Accellerase® DUET/g cellulose + 6 g Accellerase BG / kg dry matter) was similar to the glucose concentration for Accellerase® 1500 at 0.22 mL/g cellulose and higher than the glucose concentration for Accellerase® DUET at 0.15 mL/g cellulose. See FIG. 45. High
- 15 concentration of Accellerase® BG was able to reduce the viscosity of the saccharification reaction mixture. Using the enzyme composition produced from fermenting H3A/EG4 #27, at an amount of 0.15 mL/g cellulose yielded 37.5 g/kg glucose after 6-h saccharification, which was substantially higher than the glucose production for Accellerase® 1500 at 0.22 mL/g cellulose and Accellerase® DUET at 0.15 mL/g cellulose. See FIG. 45.
- 20 **[0444]** Glucose concentrations for various experimental conditions of Day 1's experiment

were measured again after 24-h saccharification. See FIG. 46. The glucose concentration and cellulose conversion were measured over time for experimental conditions 7-12 on Day 2's experiment and results are shown in FIGs. 47 and 48.

[0445] Viscosity was observed by eye on Day 1's experiment after 6-h saccharification and is summarized in Table 6. More "+" indicates less viscous saccharification reaction mixture. In general, less viscous saccharification reaction mixture (*e.g.*, thinner slurry) correlated with more glucose production.

Table 6. Viscosity observation for Day 1's experiment at 6-h

Experimental condition	Enzymes	Viscosity Observation	Glucose (g/kg)
1	Accellerase® 1500, 0.22	++	32.1
2	Accellerase® DUET, 0.15	+	27
3	Accellerase® DUET, 0.25	++++	40.8
4	Accellerase® DUET + Optimash BG	++	31.4
5	Accellerase® DUET + Optimash TBG	+	30.6
6	Accellerase® DUET + Optimash VR	+++	26.7

[0446] Viscosity of the saccharification reaction mixtures in various chambers on Day 2's experiment was observed by eye with reference to the visibility of the metal parts in each chamber. After 6-day of saccharification at 50°C, the saccharification mixture in chamber 3 (Experimental condition 9, Accellerase® DUET at 0.15 mL/g cellulose) was more viscous than the saccharification mixture in chamber 1 (Experimental condition 7) or 2 (Experimental condition 8, Accellerase® 1500 at 0.22 mL/g cellulose). Metal parts in chamber 3 could not be seen. The viscosity of the saccharification mixture in chamber 4 (Experimental condition 10, Accellerase DUET® at 0.15 mL/g cellulose + Accellerase® BG at 0.1 g/kg dry matter) was reduced compared to the viscosity of the saccharification mixture in chamber 3 (Accellerase® DUET at 0.15 mL/g cellulose). The viscosity of the saccharification mixture in chamber 5 (Experimental condition 11, Accellerase DUET® at 0.15 mL/g cellulose + Accellerase BG at 6 g/kg dry matter) was more reduced compared to the viscosity of the saccharification mixture in chamber 4 (Accellerase® DUET at 0.15 mL/g cellulose + Accellerase BG at 0.1 g/kg dry matter). Even with a high amount of Accellerase BG, the saccharification mixture (chamber 5, Accellerase DUET® at 0.15 mL/g cellulose + Accellerase BG at 6 g/kg dry matter) was still more viscous than Accellerase® 1500 at 0.22 mL/g cellulose (chambers 1 and 2). However, with the addition of the enzyme composition produced from fermenting H3A/EG4 #27, it was surprisingly found that the viscosity of the

saccharification mixture (chamber 6) was substantially reduced compared to the viscosity of the saccharification mixture in chamber 4 or 5. Metal parts in chamber 6 could be seen.

Example 21: Determining the effects of various cellulases on viscosity reduction and glucose production in saccharification process

[0447] Asaccharification process was performed by incubating Inbicon pretreated wheat straw (25% dry matter) with various enzymes in reaction chambers. The experimental conditions are shown in Table 7. In each chamber, the total mass is 10 kg. The initial pH of the wheat straw was about 3.50 and was adjusted by adding Na₂CO₃ to pH 5.0. Accellerase® 1500, Accellerase® DUET, Accellerase® BG, Optimash™ BG, and Primafast® LUNA are products available from Genecor.

Table 7

Experimental condition	Enzymes	Cellulase Loading mL/g cellulose	Viscosity Enzyme g/kg dry matter
1	Accellerase® DUET	0.15	0
2	Accellerase® 1500	0.22	0
3	Accellerase® DUET + Optimash BG	0.15	1
4	Accellerase® DUET + Optimash BG	0.15	2
5	Accellerase® DUET + Primafast LUNA	0.15	1
6	Accellerase® DUET + Primafast LUNA	0.15	2
7	Accellerase® DUET + Accellerase® BG	0.15	1
8	Accellerase® DUET + Accellerase® BG	0.15	2
9	Accellerase® DUET + Optimash BG + Accellerase® BG	0.15	1 for Optimash BG; 1 for Accellerase® BG
10	Accellerase® DUET + Accellerase® 1500	0.15 for Accellerase® DUET; 0.22 for Accellerase® 1500	0
11	H3A/Eg4#27 + Optimash BG	0.15	1
12	H3A/Eg4#27 + Optimash BG	0.15	2
13	H3A/Eg4#27 + Primafast Luna	0.15	1
14	H3A/Eg4#27 + Primafast Luna	0.15	2
15	H3A/Eg4 #27 +Accellerase® BG	0.15	1
16	H3A/Eg4#27 +Accellerase® BG	0.15	2

[0448] Glucose concentration was measured after 6 h, 24 h, 50 h, and 6 d of

saccharification. Viscosity of saccharification reaction mixture was observed by eye and measured by a viscosity meter using methods known to one skilled in the art after 6 h, 24 h, 50 h, and 6 d of saccharification.

[0449] It was found that the glucose production of each of the experimental conditions 3-16 was increased compared to the glucose production of experimental condition 1. It was further found that the viscosity of each of the experimental conditions 3-16 was reduced compared to

the viscosity of experimental condition 1.

[0450] This study also examined the glucose production and viscosity reduction in a saccharification process with the same experimental conditions as above but after a prolonged pre-hydrolysis time (such as 6 h, 9 h, 12 h, 24 h).

5 **Example 22: Ascorbic acid effect on Avicel hydrolysis by CBH1 and EG4**

[0451] Crystalline cellulose (50 μ L of 10% Avicel in 50mM Sodium Acetate, pH 5.0) reactions were initiated by mixing together combinations of purified *T. reesei* CBH1 (5 mg/g final concentration), purified *T. reesei* Eg4 (10 mg/g final concentration), ascorbic acid (50 mM stock, 8.8 g/L final concentration) and manganese solution (10 mM final concentration) as described listed in FIG. 39A. Fifty (50) mM sodium acetate buffer, pH 5.0, was added to each sample to a final volume of 300 μ L.

[0452] Reaction eppendorf tubes were vortexed and then placed in an Innova 44 incubator (New Brunswick Scientific) at 50°C, 200 rpm. Fifty (50) μ L samples were taken from each tube at three time points (2.5, 4.5, 24 h) and quenched with 50 μ L of 100 mM glycine buffer, pH 10.0. Samples were centrifuged at 3000 rpm for 5 minutes (Rotanta 460R Centrifuge, Hettich Zentrifugen) and supernatant (20 μ L) was added to 100 μ L of water in an HPLC 96-well microtiter plate (Agilent, 5042-1385). Glucose and cellobiose concentrations were measured by HPLC using Aminex HPX-87P column (300mm x 7.8mm, 125-0098) pre-fitted with guard column. The results are shown in FIG. 37.

20 [0453] Next ascorbic acid effect on Avicel hydrolysis by CBH2 and EG4 was measured. Crystalline cellulose (80 μ L of 10% Avicel in 50mM Sodium Acetate, pH 5.0) reactions were initiated by mixing together combinations of purified *T. reesei* CBH2 (5 mg/g final concentration), purified *T. reesei* Eg4 (10 mg/g final concentration), ascorbic acid (50 mM stock, 8.8 g/l final concentration) and manganese solution (10 mM final concentration) as listed in FIG. 39B. Fifty (50) mM sodium acetate buffer, pH 5.0, was added to each sample to a final volume of 500 μ L.

[0454] Reaction eppendorf tubes were vortexed and then placed in an Innova 44 incubator (New Brunswick Scientific) at 50°C, 200 rpm. Fifty (50) μ L samples were taken from each tube at three time points (5, 24, 48 h) and quenched with 50 μ L of 100 mM glycine buffer, pH 10.0. Samples were centrifuged at 3000 rpm for 5 minutes (Rotanta 460R Centrifuge, Hettich Zentrifugen) and supernatant (20 μ L) was added to 100 μ L of water in an HPLC 96-well microtiter plate (Agilent, 5042-1385). Glucose and cellobiose concentrations were

measured by HPLC using Aminex HPX-87P column (300mm x 7.8mm, 125-0098) pre-fitted with guard column. Results are shown in FIG. 38.

CLAIMS

What is claimed is:

1. A biomass saccharification mixture comprising:

a. a biomass material

b. an enzyme composition comprising a glycosyl hydrolase family 61 enzyme having endoglucanase activity, which is:

i. at least 65% in sequence identity to any one of SEQ ID NO:1-29 and 148;

ii. at least 65% in sequence identity to residues 22-344 of SEQ ID NO:27

iii. comprises at least one amino acid sequence motifs selected from the group consisting of: SEQ ID NOs: 84-91;

iv. comprises one or more sequence motifs selected from the group consisting of: (1) SEQ ID NO:84 and 88; (2) SEQ ID NOs: 85 and 88; (3) SEQ ID NO:86; (4) SEQ ID NO:87; (5) SE ID NO:84, 88 and 89; (6) SEQ ID NOs: 84, 88 and 91; (10) SEQ ID NOs: 85, 88 and 91; (11) SEQ ID NOs: 84, 88, 89 and 91; (12) SEQ ID NOs: 84, 88, 90 and 91; (13) SEQ ID NOs: 85, 88, 89 and 91; and (14) SEQ ID NOs: 85, 88, 90 and 91;

v. encoded by a polynucleotide sequence or a complement thereof that is at least 65% sequence identity to SEQ ID NO:30; or

vi. encoded by a polynucleotide sequence that hybridizes under high stringency conditions to SEQ ID NO:30 or to a complement thereof;

wherein said biomass saccharification mixture has a lower viscosity than a biomass saccharification mixture without the glycosyl hydrolyase family 61 enzyme and/or is capable of increasing the level of saccharification in the mixture as compared to the level of saccharification in a mixture having no or a lower level of glycosyl hydrolase family 61 enzyme.

2. The biomass saccharification mixture of claim 1, wherein the level of saccharification is measured by the yield of fermentable sugar after the mixture is incubated for a period of time sufficient to cause saccharification of the biomass.

3. The biomass saccharification mixture of claim 1 or 2, wherein the glycosyl hydrolase family 61 enzyme is derived from a filamentous fungus.
4. The biomass saccharification mixture of claim 3, wherein the filamentous fungus is one selected from the group: *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*,
5 *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*,
Cochliobolus, *Pyricularia*, *Chrysosporium*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*,
Aspergillus niger, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*,
10 *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*,
Bjerkandera adusta, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*,
15 *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Neurospora intermedia*, *Penicillium purpurogenum*, *Penicillium canescens*, *Penicillium solitum*, *Penicillium funiculosum* *Phanerochaete*
20 *chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*,
Trichoderma koningii, *Trichoderma longibrachiatum*, *Trichoderma reesei*,
Trichoderma viride, *Geosmithia emersonii*, or *G. stearothermophilus*.
5. The biomass saccharification mixture of any one of claims 1-4, wherein the enzyme composition further comprises at least one polypeptide having cellobiohydrolase activity and at least one polypeptide having beta-glucosidase activity.
6. The biomass saccharification mixture of claim 5, wherein the at least one polypeptide having cellobiohydrolase activity is a polypeptide encoding a *T. reesei* CBH1, Af 7A (SEQ ID NO:150), Af7B (SEQ ID NO:151), Cg7A (SEQ ID NO:152), Cg7B (SEQ ID NO:153), Tt7A (SEQ ID NO:154), Tt7B (SEQ ID NO:155), *T. reesei* CBH2, Tt6A (SEQ ID NO:156), St6A (SEQ ID NO:157), St6B (SEQ ID NO:158), or a variant thereof having at least 90% sequence identity thereto.

7. The biomass saccharification mixture of claim 5 or 6, wherein the at least one polypeptide having beta-glucosidase activity is:
- a. a polypeptide encoding an Fv3C (SEQ ID NO:100), a Pa3D (SEQ ID NO:94), an Fv3G (SEQ ID NO:96), an Fv3D (SEQ ID NO:98), a Tr3A (SEQ ID NO:102), a Tr3B (SEQ ID NO:104), a Te3A (SEQ ID NO:106), an An3A (SEQ ID NO:108), an Fo3A (SEQ ID NO:110), a Gz3A (SEQ ID NO:112), an Nh3A (SEQ ID NO:114), a Vd3A (SEQ ID NO:116), a Pa3G (SEQ ID NO:118), a Tn3B (SEQ ID NO:119), or a variant thereof having at least 90% sequence identity thereto; or
 - b. a polypeptide encoded by a polynucleotide (1) having at least 90% sequence identity to SEQ ID NO:99, 93, 95, 97, 101, 103, 105, 107, 109, 111, 113, 115, or 117; (2) hybridizes under high stringency conditions to SEQ ID NO: 99, 93, 95, 97, 101, 103, 105, 107, 109, 111, 113, 115, or 117, or to a complement thereof.
8. The biomass saccharification mixture of any one of claims 7-9, wherein the enzyme composition further comprises one or more or all of: (1) a polypeptide having xylanase activity, (2) a polypeptide having beta-xylosidase activity; and (3) a polypeptide having L-alpha-arabinofuranosidase activity.
9. The biomass saccharification mixture of claim 8, wherein the polypeptide having xylanase activity is:
- a. a polypeptide encoding a T. reesei Xyn3 (SEQ ID NO:76), T. reesei Xyn2 (SEQ ID NO:77), an AfuXyn2 (SEQ ID NO:58), and AfuXyn5 (SEQ ID NO:60), or a variant thereof having at least 90% sequence identity thereto; or
 - b. a polypeptide encoded by a polynucleotide (1) having at least 90% sequence identity to SEQ ID NO:75, 57, or 59; or (2) hybridizes under high stringency conditions to SEQ ID NO: 75, 57, or 59, or to a complement thereof.
10. The biomass saccharification mixture of claim 8, wherein the at least one polypeptide having beta-xylosidase activity is:
- a. a polypeptide encoding an Fv3A (SEQ ID NO:36), an Fv43A (SEQ ID NO:44), a Pf43A (SEQ ID NO:38), an Fv43D (SEQ ID NO:62), an Fv39A (SEQ ID NO:42), an Fv43E (SEQ ID NO:40), an Fo43A (SEQ ID NO:52), an Fv43B (SEQ ID NO:46), a Pa51A (SEQ ID NO:48), a Gz43A (SEQ ID

NO:50), a *T. reesei* Bx11 (SEQ ID NO:78), or a variant thereof having at least 90% sequence identity thereto; or

- b. a polypeptide encoded by a polynucleotide (1) having at least 90% sequence identity to SEQ ID NO:35, 43, 37, 61, 41, 39, 51, 45, 47, 49, or 159; (2)

5 hybridizes under high stringency conditions to SEQ ID NO: 35, 43, 37, 61, 41, 39, 51, 45, 47, 49, 159, or to a complement thereof.

11. The biomass saccharification mixture of claim 8, wherein the at least one polypeptide having L-alpha-arabinofuranosidase activity is:

- a. a polypeptide encoding an Af43A (SEQ ID NO:54), an Fv43B (SEQ ID NO:46), a Pf51A (SEQ ID NO:56), a Pa51A (SEQ ID NO:48), an Fv51A (SEQ ID NO:66), or a variant thereof having at least 90% sequence identity thereto; or

10

- b. a polypeptide encoded by a polynucleotide (1) having at least 90% sequence identity to SEQ ID NO:53, 45, 55, 47, or 65; (2) hybridizes under high

15

stringency conditions to SEQ ID NO: 53, 45, 55, 47, or 65, or to a complement thereof.

12. The biomass saccharification mixture of any one of claims 1-11, wherein the enzyme composition comprises (1) about 0.1 wt.% to about 50 wt.%, about 1 wt.% to about 20 wt.%, about 5 wt.% to about 15 wt.% of the polypeptide having GH61/endoglucanase activity, referencing the total weight of proteins in the enzyme composition; or (2) about 0.2 mg to about 30 mg, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 1 mg to about 5 mg of the polypeptide having GH61/endoglucanase activity per gram of cellulose, hemicelluloses or a mixture of cellulose and hemicelluloses contained in the biomass material.

20

13. The biomass saccharification mixture of any one of claims 5-12, wherein the enzyme composition comprises cellulobiohydrolase in an amount that is (1) about 0.1 wt.% to about 80 wt.%, about 5 wt.% to about 70 wt.%, about 10 wt.% to about 60 wt.%, about 20 wt.% to about 50 wt.%, or about 25 wt.% to about 50 wt.% of the total weight of proteins in the enzyme composition; or (2) about 0.2 mg to about 30 mg, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 0.5 mg to about 5 mg per gram of cellulose, hemicelluloses, or a mixture of cellulose and hemicelluloses in the biomass saccharification mixture; and comprises beta-

25

30

glucosidase in an amount that is (1) about 0.1 wt.% to about 50 wt.%, about 1 wt.% to about 30 wt.%, about 2 wt.% to about 20 wt.%, about 5 wt.% to about 20 wt.%, or about 8 wt.% to about 15 wt.% of the total weight of proteins in the enzyme composition; or (2) about 0.2 mg to about 30 mg, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 0.5 mg to about 5 mg per gram of cellulose, hemicelluloses, or a mixture of cellulose and hemicelluloses in the biomass saccharification mixture.

14. The biomass saccharification mixture of any one of claims 8-13, wherein the enzyme composition comprises (1) about 0.1 wt.% to about 50 wt.%, about 1 wt.% to about 40 wt.%, about 4 wt.% to about 30 wt.%, about 5 wt.% to about 20 wt.%, or about 8 wt.% to about 15 wt.% of the polypeptide having xylanase activity, referencing the total weight of proteins in the enzyme composition; or (2) about 0.2 mg to about 30 mg, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 0.5 mg to about 5 mg of the polypeptide having xylanase activity per gram of cellulose, hemicelluloses, or a mixture of cellulose and hemicelluloses in the biomass saccharification mixture.

15. The biomass saccharification mixture of any one of claims 8-14, wherein the enzyme composition comprises (1) about 0.1 wt.% to about 50 wt.% , about 1 wt.% to about 40 wt.%, about 2 wt.% to about 30 wt.%, about 4 wt.% to about 20 wt.%, or about 5 wt.% to about 15 wt.% of the polypeptide having beta-xylosidase activity, referencing the total weight of proteins in the enzyme composition; or (2) about 0.2 mg to about 30 mg, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 0.5 mg to about 5 mg of the polypeptide having beta-xylosidase activity per gram of cellulose, hemicelluloses, or a mixture of cellulose and hemicelluloses in the biomass saccharification mixture.

16. The biomass saccharification mixture of any one of claims 8-15, wherein the enzyme composition comprises (1) about 0.1 wt.% to about 50 wt.%, about 0.1 wt.% to about 50 wt.% , about 1 wt.% to about 40 wt.%, about 2 wt.% to about 30 wt.%, about 4 wt.% to about 20 wt.%, or about 5 wt.% to about 15 wt.% of the polypeptide having L-alpha-arabinofuranosidase activity, referencing the total weight of proteins in the enzyme composition; or (2) about 0.2 mg to about 30 mg, about 0.2 mg to about 20 mg, about 0.5 mg to about 10 mg, or about 0.5 mg to about 5 mg of the polypeptide

having L-alpha-arabinofuranosidase activity per gram of cellulose, hemicelluloses, or a mixture of cellulose and hemicelluloses in the biomass saccharification mixture.

17. The biomass saccharification mixture of any one of claims 1-16, wherein the enzyme composition is a whole cellulase composition.

5 18. The biomass saccharification mixture of claim 17, wherein the whole cellulase composition is derived from a host cell expressing a polynucleotide encoding a polypeptide having GH61/endoglucanase activity.

10 19. The biomass saccharification mixture of claim 18, wherein the polynucleotide encoding the polypeptide having GH61 family enzyme activity is heterologous to the host cell.

20. The biomass saccharification mixture of any one of claims 17-19, wherein the whole cellulase composition is derived from a host cell further expressing a polynucleotide encoding a polypeptide having cellobiohydrolase activity, and a polynucleotide encoding a polypeptide having beta-glucosidase activity.

15 21. The biomass saccharification mixture of claim 20, wherein the polynucleotide encoding the polypeptide having cellobiohydrolase activity and/or the polynucleotide encoding the polypeptide having beta-glucosidase are heterologous to the host cell.

20 22. The biomass saccharification mixture of any one of claims 17-21, wherein the whole cellulase composition is derived from a host cell expressing one or more or all of (1) a polynucleotide encoding a peptide having beta-xylosidase activity; (2) a polynucleotide encoding a polypeptide having xylanase activity; and (3) a polynucleotide peptide having L-alpha-arabinofuranosidase activity.

25 23. The biomass saccharification mixture of claim 22, wherein the polynucleotide encoding the polypeptide having beta-xylosidase activity, the polynucleotide encoding the polypeptide having xylanase activity, or the polynucleotide encoding the polypeptide having L-alpha-arabinofuranosidase activity is heterologous to the host cell.

24. The biomass saccharification mixture of any one of claims 17-23, wherein the enzyme composition is a whole broth formulation.

30 25. The biomass saccharification mixture of any one of claims 1-24, wherein one or more or all of: (1) the gene encoding the polypeptide having GH61/endoglucanase activity; (2) the gene encoding the polypeptide having cellobiohydrolase activity; (3) the gene

encoding the polypeptide having beta-glucosidase activity; (4) the gene encoding the polypeptide having beta-xylosidase activity; (5) the gene encoding the polypeptide having xylanase activity; and (6) the gene encoding the polypeptide having L-alpha-arabinofuranosidase activity are integrated into the genetic material of the host cell.

26. The biomass saccharification mixture of any one of claims 18-25, wherein the host cell is a bacterial host cell, yeast host cell, or a fungal host cell.

27. The biomass saccharification mixture of claim 26, wherein the host cell is a filamentous fungal host cell.

28. The biomass saccharification mixture of claim 27, wherein the filamentous fungal

host cell is one selected from a cell of *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *Trichoderma reesei*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Neurospora intermedia*, *Penicillium purpurogenum*, *Penicillium canescens*, *Penicillium solitum*, *Penicillium funiculosum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, or *Trichoderma viride*.

29. The biomass saccharification mixture of any one of claims 1-28, wherein the saccharification mixture is prepared by first blending the enzyme composition comprising the polypeptide having GH61/endoglucanase activity, followed by mixing the enzyme composition with the biomass.

30. The biomass saccharification mixture of any one of claims 1-29, wherein the biomass material comprises hemicellulose, cellulose, or a mixture of hemicelluloses and cellulose.
- 5 31. The biomass saccharification mixture of any one of claims 1-30, wherein the biomass material comprises glucan, xylan and/or lignin.
32. The biomass saccharification mixture of any one of claims 1-30, wherein the biomass material is selected from seeds, grains, tubers, plant waste, byproducts of food processing or industrial processing, corn cobs, corn stover, grasses, *Sorghastrum nutans*, switchgrass, perennial canes, wood, wood chips, wood processing waste, 10 sawdust, paper, paper waste, pulp, and recycled paper, potatoes, soybean, barley, rye, oats, wheat, beets, sugar cane bagasse and straw.
33. The biomass saccharification mixture of any one of claims 1-32, wherein the biomass material is subjected to pretreatment with an acid or a base.
34. The biomass saccharification mixture of claim 33, wherein the pretreated biomass is 15 adjusted to pH of about 4.0 to 6.5 before mixing with the enzyme composition.
35. The biomass saccharification mixture of any one of claims 1-34, wherein the biomass material is present in the mixture in an amount of about 5 wt.% to about 60 wt.%, about 10 wt.% to about 50 wt.%, about 15 wt.% to about 40 wt.%, about 15 wt.% to about 30 wt.%, or about 20 wt.% to about 30 wt.%, referring to the amount of 20 biomass material in its solid state relative to the total weight of the mixture.
36. A method of hydrolyzing a biomass material comprising incubating the biomass saccharification mixture of any one of claims 1-35, under conditions suitable for hydrolyzing the biomass materials in the biomass saccharification mixture and for a sufficient period of time.
- 25 37. The method of claim 36, wherein the conditions suitable for hydrolyzing the biomass materials in the biomass saccharification mixture comprises: (1) a pH of about 3.5 to about 7.0; (2) for a duration of about 2 hours or longer; and/or (3) a temperature of about 20°C to about 75°C.
38. The method of claim 36 or 37, wherein the sufficient period of time comprises a time 30 period of about 8 hours to about 72 hours.
39. The method of any one of claims 36 to 38, wherein at any given time above 2 hours, the amount of fermentable sugars is produced by the biomass saccharification mixture

is increased by at least about 5% as compared to the amount of fermentable sugars produced by a control biomass saccharification mixture comprising the same amount and type of biomass material, and the same composition of enzyme components but in the absence of the GH61/endoglucanase.

- 5 40. The method of claim 39, wherein amount of fermentable sugars produced by the biomass saccharification is increased by at least about 10%, as compared to the fermentable sugars produced by the control biomass saccharification mixture.
41. The method of any one of claims 36-40, wherein the biomass material is present in an amount of about 10 wt.% to about 50 wt.% in its solid state.
- 10 42. The method of claim 41, wherein the viscosity of the biomass saccharification mixture is reduced by at least about 5%, about 10%, about 15%, about 20%, about 25%, or more, as compared to the viscosity of the control biomass saccharification mixture comprising the same amount and type of biomass material, and the same composition of enzyme components but in the absence of the GH61/endoglucanase.
- 15 43. A method of using the composition of any one of claims 1-35 to convert a biomass material into fermentable sugars in a merchant enzyme supply model or an on-site bio-refinery model.