(12) STANDARD PATENT (11) Application No. **AU 2010333801 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Methods for improving the efficiency of simultaneous saccharification and fermentation reactions

(51) International Patent Classification(s)
C12P 7/10 (2006.01) ***C12P 19/44*** (2006.01)
C12N 1/22 (2006.01) *C12P 7/04* (2006.01)
C12N 9/24 (2006.01) *C12P 7/16* (2006.01)
C12N 9/42 (2006.01)

(21) Application No: **2010333801** (22) Date of Filing: **2010.12.17**

(87) WIPO No: **WO11/079048**

(30) Priority Data

(31) Number	(32) Date	(33) Country
61/289,917	**2009.12.23**	**US**

(43) Publication Date: **2011.06.30**
(44) Accepted Journal Date: **2015.12.17**

(71) Applicant(s)
DANISCO US INC.

(72) Inventor(s)
Gutierrez, Christina;Mitchinson, Colin;Huang, Tom T.;Diner, Bruce A.;Fagan, Paul Joseph;Hitz, William D.

(74) Agent / Attorney
Houlihan², Level 1 70 Doncaster Road, BALWYN NORTH, VIC, 3104

(56) Related Art
D2 : MAMMA ET AL: "Fungal multienzyme production on industrial by-products of the citrus-processing industry". Bioresource Technology, 2008, vol. 99, no. 7, pages 2373-2383.
D3 : US 2009/0280541 A1 (JORDAN ET AL) 12 November 2009
D1 : US 2009/0099079 A1 (EMALFARB ET AL) 16 April 2009

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 June 2011 (30.06.2011)

PCT

(10) International Publication Number
WO 2011/079048 A9

(51) International Patent Classification:
C12P 19/14 (2006.01)

(21) International Application Number:
PCT/US2010/061082

(22) International Filing Date:
17 December 2010 (17.12.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/289,917 23 December 2009 (23.12.2009) US

(71) Applicant *(for all designated States except US)*: **DANISCO US INC.** [US/US]; 925 Page Mill Road, Palo Alto, California 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **GUTIERREZ, Christina** [US/US]; Danisco Us Inc., 925 Page Mill Road, Palo Alto, California 94304 (US). **MITCHINSON, Colin** [GB/US]; Danisco Us Inc., 925 Page Mill Road, Palo Alto, California 94304 (US). **HUANG, Tom T.** [US/US]; 5322 Matthew Terrace, Fremont, California 94555 (US). **DINER, Bruce A.** [US/US]; 295 Fairville Road, Chadds Ford, Pennsylvania 19317 (US). **FAGAN, Paul Joseph** [US/US]; 10 Tenby Drive, Wilmington, Delaware 19803 (US). **HITZ, William D.** [US/US]; 404 Hillside Road, Wilmington, Delaware 19807 (US).

(74) Agent: **CHEN, Wenfang**; Danisco Us Inc., 925 Page Mill Road, Palo Alto, California 94304 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(48) Date of publication of this corrected version:
25 August 2011

(15) Information about Correction:
see Notice of 25 August 2011

(54) Title: METHODS FOR IMPROVING THE EFFICIENCY OF SIMULTANEOUS SACCHARIFICATION AND FERMENTATION REACTIONS

Primer sequences for the construction of *bxl1* deletion cassette

Primer Name	Seq ID No.	Primer sequence (5'→3')	Description
MH290	27	caaGGCGCGCCaagtATAACTTCGTATAAT GTATGCTATACGAAGTTATCGGCCGGCG TATTGGGTG⁻TACG	hph reverse primer AscI loxP site
MH292	28	GAAGGCGCGCCACAGATAACTTCGTA⁻A GCATACATTATACGAAGTTATcctgggcttgtg actggtcgcgag	hph forward primer full promoter AscI loxP site
MH375	29	ccatgtcacctgtcttgaacac	Bxl1 5' forward
MH376	30	**caaggcgc**GCCATCTCTT⁻CGATCTCAACA G	Bxl1 5' reverse AscI
MH377	31	gattgcgatcgccgtctacaacgtttcaacc	Bxl1 3' forward AsiSI AclI
MH378	32	GGTCCAACCTTGAATGTAACAGC	Bxl1 3' reverse primer
MH379	33	gtgtcgctgaacataaggtctc	Bxl1 deletion nested forward primer
MH380	34	CCTCCATTCTTCCAACAAGCC	Bxl1 deletion nested reverse primer

TABLE 4

(57) Abstract: The present disclosure is directed, in a first aspect, to the use of inverting beta-xylosidase enzymes to reduce byproduct formation and increase the yield of fermentation products, as well as, in a second aspect, to the use of retaining beta-xylosidase enzymes to improve production of alkyl-beta-xylopyranoside compounds, in a simultaneous saccharification and fermentation reactions.

WO 2011/079048 A9

METHODS FOR IMPROVING THE EFFICIENCY OF SIMULTANEOUS SACCHARIFICATION AND FERMENTATION REACTIONS

1. CROSS-REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application No. 61/289,917, filed December 23, 2009, which is incorporated by reference in its entirety.

2. FIELD OF THE DISCLOSURE

The present disclosure is generally directed to methods and compositions for improving product yield from simultaneous saccharification and fermentation reactions.

3. BACKGROUND

Bioconversion of renewable lignocellulosic biomass to a fermentable sugar that is subsequently fermented to produce an alcohol (*e.g.*, ethanol or "bioethanol"), which can serve as an alternative to liquid fuels, has attracted intensive attention of researchers since the 1970s, when the oil crisis occurred because OPEC decreased the output of petroleum (Bungay, "Energy: the biomass options". NY: Wiley; 1981; Olsson and Hahn-Hagerdal, 1996, Enzyme Microb. Technol. 18:312-31; Zaldivar *et al.*, 2001, Appl. Microbiol. Biotechnol. 56:17-34; Galbe and Zacchi, 2002, Appl. Microbiol. Biotechnol. 59:618-28). Ethanol has been widely used as a 10% blend to gasoline in the USA or as a neat fuel for vehicles in Brazil in the last two decades. The importance of fuel bioethanol will increase in parallel with skyrocketing prices for oil and gradual depletion of its sources.

Lignocellulosic biomass is predicted to be a low-cost renewable resource that can support the sustainable production of biofuels (*e.g.*, bioethanol) on a large enough scale to significantly address the world's increasing energy needs. Lignocelllulosic materials include, without limitation, corn stover (the corn plant minus the kernels and the roots), forestry residues such as sawdust and paper, yard waste from municipal solid waste, herbaceous plants such as switchgrass, and woody plants such as poplar trees. Lignocellulosic biomass has three major components: hemicellulose, cellulose, and lignin. Hemicellulose is an amorphous, branched polymer that is usually composed primarily of five sugars (arabinose, galactose, glucose, mannose, and xylose). Cellulose is a large, linear polymer of glucose molecules typically joined together in a highly crystalline structure due to hydrogen bonding between parallel chains. Lignin is a complex phenyl-propane polymer.

The biological processing of lignocellulosic biomass involves using cellulases and hemicellulases to release sugars from hemicellulose and cellulose, respectively, typically

by hydrolysis reactions. The resulting sugars are then fermented into biofuels such as bioethanol using suitable fermenting microorganisms.

The glucose released when cellulose is broken down by cellulases can often be a potent inhibitor of this class of enzymes. To reduce glucose accumulation during cellulose breakdown (or "saccharification" herein), a fermenting microorganism can be added to convert the released sugars into bioethanol at the same time the sugars are revealed from saccharification. This configuration is called simultaneous saccharification and fermentation ("SSF"). Generally, SSF offers better/higher rates, yields, and concentrations of ethanol produced than a separate hydrolysis and fermentation ("SHF") configuration, despite operating at lower temperatures than are optimal for most enzymes involved in these fermentation processes. Nonetheless, the typical SSF reaction can be exceedingly lengthy, lasting, for example, several days in order to achieve modest ethanol concentrations (*see, e.g.*, Kadam *et al.*, 2004, Biotechnol. Progr. 20(3):705).

Accordingly, there exists a need in the art to identify methods and compositions related thereto for improving the efficiency of SSF reactions and increasing the yield of biofuels such as bioethanol.

4. SUMMARY

The instant disclosure is based on the discovery that the presence of certain β-xylosidases in simultaneous saccharification and fermentation ("SSF") reactions results in a rapid accumulation of alkyl-β-xylopyranoside byproducts. In particular, it is discovered that certain β-xylosidases with a retaining mechanism of action, when used in SSF reactions, can result in rapid accumulation of alkyl-β-xylopyranoside byproducts that would lead to a reduced yield of fermentation products. The present disclosure is further based on the discovery that the presence of certain other β-xylosidases with an inverting mechanism of action in SSF reactions can reduce or minimize the accumulation of alkyl-β-xylopyranoside byproducts. The inclusion of β-xylosidases with an inverting mechanism of action in SSF reactions has been found to improve the yield of fermentation products.

For the purpose of this disclosure, a β-xylosidase with an inverting mechanism of action is also referred to as "an enzyme with inverting β-xylosidase activity," "an inverting β-xylosidase," or "an inverting β-xylosidase polypeptide." Conversely, a β-xylosidase with a retaining mechanism of action is also referred to as "an enzyme with retaining β-xylosidase activity," "a retaining β-xylosidase," or "a retaining β-xylosidase polypeptide."

Accordingly, provided herein are improved methods for conducting SSF reactions that entail reducing the amount of retaining β-xylosidases. Provided herein are also improved methods for conducting SSF reactions that entail increasing the amount of inverting β-xylosidases. Provided further herein are improved methods for conducting SSF reaction that entail decreased amount of retaining β-xylosidases and increased amount of inverting β-xylosidases. Compositions related to the above-described improved methods and other methods described herein are also comtemplated.

Accordingly, there is provided a method for simultaneous saccharification and fermentation (SSF) comprising culturing a complete fermentation medium, said complete fermentation medium comprising at least one fermenting microorganism, at least one xylan-containing biomass, at least one cellulase, at least one hemicellulase, at least one retaining β-xylosidase, and at least one inverting β-xylosidase, for a period and under conditions suitable for producing a fermentation product, wherein the inverting β-xylosidase is an Fv43D, a Pf43A, an Fv43E, an Fv43B, an Af43A, an Fo43A, a Gz43A, or a XynB3 polypeptide, and the complete fermentation medium comprises a greater amount of inverting b-xylosidases than that of retaining b-xylosidases on a mole basis, molecular weight basis, or on both a mole basis and a molecular weight basis.

In certain aspects, the present invention provides SSF methods comprising culturing a complete fermentation medium, said complete fermentation medium comprising at least one fermenting microorganism, at least one xylan-containing biomass, at least one cellulase, at least one hemicellulase, and at least one enzyme with inverting β-xylosidase activity, for a period and under conditions suitable for formation of a fermentation product.

In the present disclosure, the one or more enzymes (or alternatively and interchangeably stated as "at least one enzyme" herein) with inverting β-xylosidase activity can be present in said complete fermentation medium in an amount effective to reduce short chain alkyl-β-xylopyranoside ("AXP") (*e.g.*, methyl-β-xylopyranoside ("MXP"), ethyl-β-xylopyranoside ("EXP"), propyl-β-xylopyranoside ("PXP"), or butyl-β-xylopyranoside ("BXP")) formation, as compared to a control fermentation medium lacking said one or more enzymes with inverting β-xylosidase activity. For example, such enzyme(s) are present in an amount effective to (a) reduce the amount of AXP formation by at least 20%, by at least 30%, by at least 40%, by at least 50%, by at least 60%, by at least 70%, or by at least 80%, and/or (b) to increase the yield (*e.g.*, by at least 0.1%, at least 0.5%, at least 0.7%, at least 1%, at least 2%, at least 3%, at least 5%, at least 7.5%, or at least 10%) of the fermentation product (*e.g.*, an alcohol, such as, but not limited to, methanol, ethanol, propanol, propane-1,3-diol, or butanol), as compared to that of a

control fermentation medium lacking said one or more enzymes with inverting β-xylosidase activity. In certain embodiments, such enzyme(s) are present in an amount effective to reduce the amount of AXP formation to a level that is within 50%, within 40%, within 30%, within 20%, or within 10% above the level of AXP at the reaction equilibrium, as compared to that of a control fermentation medium lacking said one or more enzymes with inverting β-xylosidase activity.

In certain aspects, the fermenting microorganism is capable of producing an alcohol, for example, methanol, ethanol, propanol, propane-1,3-diol, or butanol, from at least one fermentable carbon source. In certain aspects, the fermenting microorganism

is a bacterium such as a *Zymomonas mobilis* or a fungus such as a yeast or a filamentous fungus.

In certain aspects, the at least one inverting β-xylosidase is a GH43 family enzyme. In certain embodiments, the inverting β-xylosidase is selected from an Fv43D, a Pf43A, an Fv43E, an Fv43B, an Af43A, an Fo43A, a Gz43A, or a XynB3 polypeptide. Specifically, the Fv43D polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:2, or to residues 21 to 350 of SEQ ID NO:2. The Pf43A polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:8, or to residues 21 to 445 of SEQ ID NO:8. The Fv43E polypepide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:10 or to residues 19 to 530 of SEQ ID NO:10. The Fv43B polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:12, or to residues 17 to 574 of SEQ ID NO:12. The Af43A polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:14 or to residues 15-558 of SEQ ID NO:14. The Fo43A polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:24, or to residues 21 to 348 of SEQ ID NO:24. The Gz43A polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at

least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:22, or to residues 19 to 340 of SEQ ID NO:22. The XynB3 polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:25.

In certain aspects, the amount of inverting β-xylosidase polypeptides in said complete medium is at least about 0.2 mg, at least about 0.3 mg, at least about 0.4 mg, at least about 0.5 mg, at least about 0.7 mg, at least about 1 mg, at least about 2 mg, or at least about 3 mg per gram of xylan in said xylan-containing biomass, which is also a component of the complete fermentation medium. In other aspects, the amount of inverting β-xylosidase polypeptides is about 3 mg or less, about 2 mg or less, about 1.5 mg or less, about 1.0 mg or less, about 0.7 mg or less, about 0.5 mg or less, about 0.4 mg or less, about 0.3 mg or less, or about 0.2 mg or less per gram of xylan in said xylan-containing biomass. In certain aspects, the amount of inverting β-xylosidase polypeptides in said complete medium ranges from, for example, (a) 0.4 mg to 10 mg, (b) 0.5 mg to 2 mg, (c) 0.4 mg to 5 mg, (d) 0.5 mg to 1.5 mg, (e) 1 mg to 2 mg, (f) 0.3 mg to 3 mg, (g) 0.2 mg to 5 mg, (h) 0.3 mg to 5 mg, or (i) 0.3 mg to 10 mg, per gram of xylan in said xylan-containing biomass, or the amount is within a range whose upper and lower limits are each independently selected from the foregoing values.

In certain aspects, the amount of inverting β-xylosidase polypeptide(s) in said complete fermentation medium exceeds the amount of retaining β-xylosidase polypeptide(s), on a mole basis, on a molecular weight basis, or on both a mole basis and a molecular weight basis. In specific embodiments, the ratio of inverting β-xylosidase polypeptide(s) to retaining β-xylosidase polypeptide(s) is at least 2:1, at least 3:1, at least 4:1, at least 5:1, at least 10:1, or at least 50:1, on a mole basis, on a molecular weight basis, or on both a mole basis and a molecular weight basis. In specific embodiments, enzyme(s) with retaining β-xylosidase activity are absent from or undetectable in the complete fermentation medium. In certain embodiments, there is no detectable retaining β-xylosidase activity in the complete fermentation medium.

According to the methods described herein, the culturing of the complete fermentation medium is conducted under continuous, batch, or fed-batch conditions. For example, the culturing of the complete fermentation medium of the invention is a continuous SSF reaction, a batch-type SSF reaction, or a fed-batch type SSF reaction.

The methods of the present disclosure, in certain aspects, further encompass the formation of the complete fermentation medium prior to the culturing step. For example, the complete fermentation medium can be formed by combining (a) at least one fermenting microorganisms, (b) at least one xylan-containing biomass, (c) at least one cellulase, (d) at least one hemicellulase, (e) at least one inverting β-xylosidase, and (f) a medium lacking one or more of the components (a) to (e). In specific embodiments, the at least one cellulase can be present in the form of a cellulase preparation. For example the cellulase preparation can be a whole cellulase preparation, which can optionally also include the at least one hemicellulase. In specific embodiments, the cellulase preparation is a culture broth from a filamentous fungal culture, *e.g.*, a *T. reesei* culture prepared using a *T. reesei* cell. In a certain aspect, the *T. reesei* cell has been engineered such that the native β-xylosidase gene is inactivated or deleted. It should be noted that a "*T. reesei* cell [that] has been engineered such that the native β-xylosidase gene is inactivated or deleted" includes not only the original or parental cell, in which the inactivation first took place, but also progeny of that cell wherein the native β-xylosidase gene is inactivated or deleted.

In certain aspects, the methods of the present disclosure pertain to culturing a fermentation broth comprising at least one xylan-containing biomass. In certain aspects, the xylan-containing biomass is, for example, corn stover, bagasse, sorghum, giant reed, elephant grass, miscanthus, Japanese cedar, wheat straw, switchgrass, hardwood pulp, or softwood pulp. For example, the xylan-containing biomass can suitably be added to the SSF reaction in the form of a slurry. For example, the xylan-containing biomass can be added to the SSF reaction in the form of a solid. Accordingly, the xylan-containing biomass can suitably be added to the SSF reaction in either a liquid form (which can be, for example, a solution, a suspension, or a mixture of solids and liquid) or in a solid form. In certain embodiments, the xylan-containing biomass has been subject to pre-treatment.

After the SSF reaction has taken place, optionally to completion, a recovery step can follow, wherein the fermentation product (*e.g.*, ethanol, methanol, propanol, propane-1,3-diol, or butanol) is recovered.

The present disclosure further provides a complete fermentation medium suitable for use in the present methods, for example as described hereinabove and hereinbelow with respect to cellulase, hemicellulase, β-xylosidase and fermenting microorganism components.

The present disclosure also provides a *T. reesei* cell that has been engineered such that the native β-xylosidase gene is inactivated or deleted. The *T. reesei* cell can

be engineered to recombinantly express an enzyme of the GH43 family, for example, an enzyme selected from an Fv43D, a Pf43A, an Fv43E, an Fv43B, an Af43A, an Fo43A, a Gz43A, or a XynB3 polypeptide. For example, the Fv43D polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:2 or to residues 21 to 350 of SEQ ID NO:2. The Pf43A polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:8, or to residues 21 to 445 of SEQ ID NO:8. The Fv43E polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:10, or to residues 19 to 530 of SEQ ID NO:10. The Fv43B polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:12 or to residues 17 to 574 of SEQ ID NO:12. The Af43A polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:14, or to residues 15-558 of SEQ ID NO:14. The Fo43A polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:24 or to residues 21 to 348 of SEQ ID NO:24. The Gz43A polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:22, or to residues 19 to 340 of SEQ ID NO:22. The XynB3 polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%,

at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:25.

The present disclosure further provides methods for producing cellulase polypeptides, comprising (a) culturing such a *T. reesei* cell, such as one described herein above and/or herein below, under conditions that allow the production of one or more cellulase polypeptides, and (b) recovering the cellulase polypeptides, for example, by recovering a culture broth comprising the cellulase polypeptides. The present disclosure further provides a culture broth produced by the *T. reesei* cell, and its use in saccharification reactions, including, for example, SSF reactions.

In certain aspects, the present disclosure provides a complete fermentation medium composition, comprising, at least one inverting β-xylosidase, at least one fermenting microorganism, at least one xylan-containing biomass, at least one cellulase, at least one hemicellulase, and a fermentation medium. The inverting β-xylosidase is, for example, a GH43 family enzyme. The inverting β-xylosidase can, in certain aspects, be one selected from an Fv43D, a Pf43A, an Fv43E, an Fv43B, an Af43A, an Fo43A, a Gz43A, or a XynB3 polypeptide. The fermenting microorganism is suitably one that is capable of fermenting a suitable carbon source such as a xylan-containing biomass, or sugars, such as glucose, xylose, arabinose, mannose, galactose, or oligosaccharides, directly or indirectly into a desired fermentation product, including, for example, methanol ("MeOH"), ethanol ("EtOH"), propanol, propane-1,3-diol, or butanol. The fermenting microorganism can be selected from a fungus, such as, for example, a yeast or a filamentous fungus, a bacterium, such as a *Zymomonas mobillis* or a *Clostridium thermocellum*. Suitable carbon sources or substrates include, for example, xylan-containing biomass substrates, selected from, for example, lingocellulosic substrates or other carbohydrate-containing raw materials. Certain of the lignocellulosic substrates can, for example, comprise cellulose, hemicellulose, and/or lignin.

The cellulase is, for example, a β-glucosidase, an endoglucanase, or a cellobiohydrolase polypeptide. Enzymes are referred to herein either by their names, or by the enzyme families to which they belong (*e.g.*, the GH43 family enzyme; or enzyme classified in or under EC 3.2.1.91); when they are referred to by their names, they can also be referred to interchangeably as a "[name] polypeptide (*e.g.*, a β-glucosidase can also be referred to, interchangeably, as a β-glucosidase polypeptide). The cellulase can also be, for example, in the form of a whole cellulase preparation. The hemicellulase is, for example, a xylanase, a β-xylosidase, an L-α-arabinofuranosidase, or an accessory

protein. The whole cellulase preparation can comprise one or more of the hemicellulase polypeptides in certain embodiments.

The fermentation medium can be one that results from a partial saccharification process, or one that comprises certain amounts of the products of saccharification. Such a composition can be suitably used in a saccharification reaction, including, for example, an SSF reaction, under conditions that allow production of the fermentation product(s) of interest.

In certain aspects, the one or more cellulases and/or the one or more hemicellulases are produced by a genetically engineered microorganism wherein the gene encoding the one or more (if more than one native β-xylosidase is present) native β-xylosidases have been deleted or there is no detectable native β-xylosidase activity. In certain embodiments, the microorganism engineered to produce the one or more cellulases and/or one or more hemicellulases does not comprise a retaining β-xylosidase or has no detectable retaining β-xylosidase activity.

In some aspects, the instant disclosure pertains to an improved method for conducting an SSF reaction on a xylan-containing biomass, in order to obtain a bioethanol fermentation product, wherein the method comprises culturing a fermentation medium comprising at least one cellulase, at least one hemicellulase, and at least one fermenting microorganism, wherein the improvement comprises the use of an enzyme with inverting β-xylosidase activity.

In related aspects, the instant disclosure also provides methods of improving production of alkyl-β-xylopyranosides, which are known to be useful and valuable for a number of industrial applications, in situations where such production is desirable. For example, alkyl-β-xylopyranoside can suitably be used as chemical intermediates in the synthesis of alkyl-glucosides, which is useful as biodegradable surfactants and emulsifiers (*see, e.g.,* K. Schmid & H. Tesmann, 2001, Alkyl Polyglucosides, in Detergency of Specialty Surfactants, Surfactant science series, vol. 98; (F.E. Fried ed.); Marcel Dekker Inc., NY, pp.1-70). These compounds are also inducers themselves or can be used to prepare inducers of xylanase production in a number of microorganisms (*see, e.g.,* M. Marui *et al.*, 1985, Agric. Biol. Chem. 49(12):3399-3407; H. Shinoyama *et al.,* 1988, Agric. Biol. Chem. 52(9): 2197-2202). Various alkyl-pyranosides can, in addition, be used as primers for chondroitin sulphate and stimulants of the biosynthesis of proteoglycans (*see, e.g.,* H. Shinoyama *et al.,* 1988, Agric. Biol. Chem. 52(9): 2197-2202). The inclusion or increased production of β-xylosidases with a retaining

mechanism of action in SSF reactions can be used to improve the yield of these useful alkyl-β-xylopyranosides in another aspect of the present invention.

Accordingly, in certain embodiments, the method of the disclosure comprises increasing an amount of retaining β-xylosidases. Provided herein are also improved methods of conducting SSF reactions that entail increasing the amount of retaining β-xylosidases. In a further example, an improved method of conducting SSF reactions that entails increasing the amount of retaining β-xylosidases while decreasing the amount of inverting β-xylosidases is contemplated.

In other aspects, the present invention provides SSF methods comprising culturing a complete fermentation medium, said complete fermentation medium comprising at least one fermenting microorganism, at least one xylan-containing biomass, at least one cellulase, at least one hemicellulase, and at least one enzyme with retaining β-xylosidase activity, for a period and under conditions suitable for formation of of alkyl-β-xylopyranoside, such as, for example, methyl-β-xylopyranoside ("MXP"), ethyl-β-xylopyranoside ("EXP"), propyl-β-xylopyranoside ("PXP"), or butyl-β-xylopyranoside ("BXP").

In certain aspects, the at least one enzyme with retaining β-xylosidase activity can be present in said complete fermentation medium in an amount effective to increase short chain alkyl-β-xylopyranoside ("AXP") (*e.g.*, methyl-β-xylopyranoside ("MXP"), ethyl-β-xylopyranoside ("EXP"), propyl-β-xylopyranoside ("PXP"), or butyl-β-xylopyranoside ("BXP")) formation, as compared to a control fermentation medium lacking or having lesser amount of said enzymes with retaining β-xylosidase activity. For example, such enzyme(s) are present in an amount effective to increase the amount of AXP formation by at least 20%, by at least 30%, by at least 40%, by at least 50%, by at least 60%, by at least 70%, or by at least 80%, as compared to that of a control fermentation medium lacking or having a lesser amount of said one or more enzymes with retaining β-xylosidase activity.

In certain aspects, the fermenting microorganism is capable of producing a number of short chain alkyl-β-xylopyranoside ("AXP") compounds, including, without limitation, methyl-β-xylopyranoside ("MXP"), ethyl-β-xylopyranoside ("EXP"), propyl-β-xylopyranoside ("PXP"), or butyl-β-xylopyranoside ("BXP") compounds. In some aspects the fermenting microorganism is a bacterium such as a *Zymomonas mobilis* or a fungus such as a yeast or a filamentous fungus.

In certain aspects, the at least one retaining β-xylosidase is a GH3, GH30, GH31, GH39, GH52, GH54, or GH116 family enzyme. In certain embodiments, the retaining

β-xylosidase is selected from a XlnD, an Fv30A, an Fv30B, an Fv39A, an Fv39B, a XynB, a XylA, or a Xyl1 polypeptide. Specifically, the XlnD polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:40, or to residues 18-804 of SEQ ID NO:40. The Fv30A polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:42, or to residues 20-537 of SEQ ID NO:42. The Fv30B polypepide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:44 or to residues 25-485 of SEQ ID NO:44. The Fv39A polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:46, or to residues 20-439 of SEQ ID NO:46. The Fv39B polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:48 or to residues 19-456 of SEQ ID NO:48. The XynB polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:50. The XylA, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:52. The Xyl1 polypeptide, if present in the complete fermentation medium, is a polypeptide comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at

least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:54, or to residues 22-500 of SEQ ID NO:54.

In certain aspects, the amount of retaining β-xylosidase polypeptides in said complete medium is at least about 0.2 mg, at least about 0.5 mg, at least about 0.7 mg, at least about 1 mg, at least about 2 mg, or at least about 5 mg per gram of xylan in said xylan-containing biomass, which is also a component of the complete fermentation medium. In other aspects, the amount of inverting β-xylosidase polypeptide(s) is about 10 mg or less, about 5 mg or less, about 2 mg or less, about 1 mg or less, about 0.7 mg or less, about 0.5 mg or less, or about 0.2 mg or less per gram of xylan in said xylan-containing biomass. In certain aspects, the amount of retaining β-xylosidase polypeptides in said complete medium ranges from, for example, (a) 0.2 mg to 10 mg, (b) 0.2 mg to 5 mg, (c) 0.5 mg to 5 mg, (d) 1 mg to 10 mg, (e) 2 mg to 10 mg, (f) 0.2 to 5 mg, (g) 0.2 mg to 2 mg, or (h) 0.5 mg to 10 mg, per gram of xylan in said xylan-containing biomass, or the amount is within a range whose upper and lower limits are each independently selected from the foregoing values.

In certain aspects, the amount of retaining β-xylosidase polypeptide(s) in said complete fermentation medium exceeds the amount of inverting β-xylosidase polypeptide(s), on a mole basis, on a molecular weight basis, or on both a mole basis and a molecular weight basis. In specific embodiments, the ratio of retaining β-xylosidase polypeptides to inverting β-xylosidase polypeptides is at least 2:1, at least 3:1, at least 4:1, at least 5:1, at least 10:1, or at least 50:1, on a mole basis, on a molecular weight basis, or on both a mole basis and a molecular weight basis. In specific embodiments, enzymes with inverting β-xylosidase activity are absent from or undetectable in the complete fermentation medium. In certain embodiments, there is no detectable inverting β-xylosidase activity in the complete fermentation medium.

According to the method describe herein, the culturing of the complete fermentation medium is conducted under continuous, batch, or fed-batch conditions. For example, the culturing of the complete fermentation medium of the invention is a continuous SSF reaction, a batch-type SSF reaction, or a fed-batch type SSF reaction.

The methods of the present disclosure, in certain aspects, further encompass the formation of the complete fermentation medium prior to the culturing step. For example, the complete fermentation medium can be formed by combining (a) at least one fermenting microorganisms, (b) at least one xylan-containing biomass, (c) at least one cellulase, (d) at least one hemicellulase, (e) at least one retaining β-xylosidase, and (f) a

medium lacking one or more of the components (a) to (e). In specific embodiments, the at least one cellulase can be present in the form of a cellulase preparation. For example the cellulase preparation can be a whole cellulase preparation, which can optionally also include the at least one hemicellulase. In specific embodiments, the cellulase preparation is a culture broth from a filamentous fungal culture, *e.g.*, a *T. reesei* culture prepared using a *T. reesei* cell. In a certain aspect, the *T. reesei* cell has been engineered such that either the native retaining β-xylosidase gene is overexpressed or that a foreign retaining β-xylosidase gene is introduced and expressed therein. It should be noted that a "*T. reesei* cell [that] has been engineered such that either the native retaining β-xylosidase gene is overexpressed or that a foreign retaining β-xylosidase gene is introduced and expressed therein" includes not only the original or parental cell, in which the inactivation first took place, but also progeny of that cell.

In certain aspects, the methods of the present disclosure pertain to culturing a fermentation broth comprising at least one xylan-containing biomass. In certain aspects, the xylan-containing biomass is, for example, corn stover, bagasse, sorghum, giant reed, elephant grass, miscanthus, Japanese cedar, wheat straw, switchgrass, hardwood pulp, or softwood pulp. For example, the xylan-containing biomass can suitably be added to the SSF reaction in the form of a slurry. For example, the xylan-containing biomass can be added to the SSF reaction in the form of a solid. Accordingly, the xylan-containing biomass can suitably be added to the SSF reaction in either a liquid form (which can be, for example, a solution, a suspension, or a mixture of solids and liquid) or in a solid form. In certain embodiments, the xylan-containing biomass has been subject to pre-treatment.

After the SSF reaction has taken place, optionally to completion, a recovery step can follow, wherein the AXP product (*e.g.*, MXP, EXP, PXP, or BXP) is recovered.

The present disclosure further provides a complete fermentation medium suitable for use in the present methods, for example as described hereinabove and hereinbelow with respect to cellulase, hemicellulase, β-xylosidase and fermenting microorganism components.

The present disclosure also provides a *T. reesei* cell that has been engineered such that the native retaining β-xylosidase gene is overexpressed, or that a foreign retaining β-xylosidase gene is expressed therein. The *T. reesei* cell can be engineered to recombinantly express an enzyme of the GH3, GH30, GH31, GH39, GH52, GH54, or GH116 family, for example, one selected from a XlnD, an Fv30A, an Fv30B, an Fv39A, an Fv39B, a XynB, a XylA, or a Xyl1 polypeptide. For example, the XlnD polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at

least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:40 or to residues 18-804 of SEQ ID NO:40. The Fv30A polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:42, or to residues 20-537 of SEQ ID NO:42. The Fv30B polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:44, or to residues 25-485 of SEQ ID NO:44. The Fv39A polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:46 or to residues 20-439 of SEQ ID NO:46. The Fv39B polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:48, or to residues 19-456 of SEQ ID NO:48 The XynB polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:50. The XylA polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:52, or to residues 19-705 of SEQ ID NO:52. The Xyl1 polypeptide has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to SEQ ID NO:54, or to residues 22-500 of SEQ ID NO:54.

The present disclosure further provides methods for producing cellulase polypeptides, comprising (a) culturing such a *T. reesei* cell, such as one described herein above and/or herein below, under conditions that allow the production of one or more

cellulase polypeptides and (b) recovering the cellulase polypeptides, for example, by recovering a culture broth comprising the cellulase polypeptides. The present disclosure further provides a culture broth produced by the *T. reesei* cell, and its use in saccharification reactions, including, for example, SSF reactions.

In certain aspects, the present disclosure provides a complete fermentation medium composition, comprising, at least one retaining β-xylosidase, at least one fermenting microorganism, at least one xylan-containing biomass, at least one cellulase, at least one hemicellulase, and a fermentation medium. The retaining β-xylosidase is, for example, either a native retaining β-xylosidase that is overexpressed or a foreign β-xylosidase that is introduced into a suitable host cell. The retaining β-xylosidase is, for example, a GH3, GH30, GH31, GH39, GH52, GH54, or GH 116 family enzyme. The retaining β-xylosidase can, in certain aspects, be one selected from a XlnD, an Fv30A, an Fv30B, an Fv39A, an Fv39B, a XynB, a XylA, or a Xyl1 polypeptide. The fermenting microorganism is suitably one that is capable of fermenting a suitable carbon source such as a xylan-containing biomass, or sugars, such as glucose, xylose, arabinose, mannose, galactose, or oligosaccharides, directly or indirectly into a desired fermentation product, including, for example, methanol, ethanol, propanol, propane-1,3-diol, or butanol. The fermenting microorganism can be a fungus, such as, for example, a yeast or a filamentous fungus, or a bacterium, such as a *Zymomonas mobillis* or a *Clostridium thermocellum*. Suitable carbon sources or substrates can be, for example, xylan-containing biomass substrates, selected from, for example, lingocellulosic substrates or other carbohydrate-containing raw materials. Certain of the lignocellulosic substrates can, for example, comprise cellulose, hemicellulose, and/or lignin.

The cellulase is, for example, a β-glucosidase, an endoglucanase, or a cellobiohydrolase polypeptide. The cellulase can also be, for example, in the form of a whole cellulase preparation. The hemicellulase is, for example, a xylanase, a β-xylosidase, an L-α-arabinofuranosidase, or an accessory protein. The whole cellulase preparation can comprise one or more hemicellulase polypeptides in certain embodiments.

The fermentation medium can be one that results from a partial saccharification process, or one that comprises certain amounts of the products of saccharification. Such a composition can be suitably used in a saccharification reaction, including, for example, an SSF reaction, under conditions that allow production of the AXP compounds, including, for example, MXP, EXP, PXP, or BXP.

In certain aspects, the one or more cellulases or the one or more hemicellulases are produced by a genetically engineered microorganism wherein the gene encoding the one or more (if more than one native β-xylosidase is present) native β-xylosidases have been overexpressed or a gene encoding a foreign β-xylosidase has been introduced and/or expressed. In certain embodiments, the microorganism engineered to produce one or more cellulases and/or one or more hemicellulases has an increased expression of retaining β-xylosidase activity over a control microorganism which did not undergo the same genetic engineering. In certain embodiments, the microorganism engineered to produce one or more cellulases and/or one or more hemicellulases does not comprises an inverting β-xylosidase or has no detectable inverting β-xylosidase activity.

In certain related aspects, the instant disclosure pertains to an improved method for conducting an SSF reaction on a xylan-containing biomass, in order to obtain an AXP product wherein the method comprises culturing a fermentation medium comprising at least one cellulase, at least one hemicellulase, and at least one fermenting microorganism, wherein the improvement comprises the use of a cellulase preparation made from a *T. reesei* cell, which has been engineered to overexpress the native β-xylosidase gene, or to express a foreign β-xylosidase. In some embodiments, the native β-xylosidase gene that is overexpressed or the foreign β-xylosidase gene that is expressed is a gene encoding a retaining β-xylosidase.

All publications, patents, patent applications, GenBank sequences and ATCC deposits cited herein are hereby expressly incorporated by reference for all purposes.

5. BRIEF DESCRIPTION OF THE FIGURES AND TABLES

Table 1: EXP formation with recombinant *Zymomonas mobilis* under the SSF conditions from cob and glucose/xylose.

Table 2: EXP formation using various types of biomass substrates.

Table 3: Time course of EXP and/or xylose formation (expressed as RI area, proportional to mg/mL) from xylobiose (20 mg/mL) in 50 mM NaCitrate, pH 4.7 plus 0.9 M EtOH at 46 ºC in the presence of Multifect® Xylanase ("MF," 560 µg/mL) and purified Fv43D (36 µg/mL) or Fv3A (54 µg/mL).

Table 4: Primer sequences for construction of *bxl1* deletion cassette.

Table 5: Primer sequences for construction of *F. verticillioides* β-xylosidase Fv43D expression cassette.

Table 6: provides a summary of the sequence identifiers for certain enzymes used in SSF reactions.

Figure 1: HPLC chromatograms of samples taken after 48 hrs of incubation at 46 ºC of xylose (10 mg/mL) in 50 mM sodium citrate, pH 4.6, Multifect® Xylanase (1.35 mg/mL) and with no alcohol, with ethanol ("EtOH"), with methanol ("MeOH"), or with n-propanol ("n-PrOH"), each at 0.72 M. HPLC conditions: column HPX-87H at 60 ºC, 0.60 mL/min 0.01 N H_2SO_4, RI detector.

Figure 2: Time course following the appearance/formation of alkyl-xylopyranosides ("AXP") under the the same conditions as those described in the experiments of Figure 1. The amounts of fermentation products formed are expressed as ratios of the areas of the product peaks compared to those of the xylose peaks.

Figure 3: NMR spectrum indicating the presence of ethyl-β-xylopyranoside.

Figure 4: EXP formation with yeast and wild-type *Zymomonas mobilis* under SSF conditions.

Figure 5: EXP formation under the yeast SSF conditions with and without the addition of *T. reesei* Bxl1.

Figure 6: EXP dose response following addition of *T. reesei* Bxl1 to the enzyme complex produced from integrated *T. reesei* strain #229 under the recombinant *Zymomonas* SSF conditions.

Figure 7: EXP formation during an SSF reaction with the following enzyme configurations/mixtures for saccharification: Accellerase™ 1500 ("A1500")+Multifect® Xylanase, Accellerase™ 1500 ("A1500") +XlnA, and an enzyme complex produced from the integrated *T. reesei* strain #229 with the addition of a hemicellulase (Fv3A, Fv51A, or Bxl1).

Figure 8: EXP formation during an SSF reaction using various purified cellulase enzymes and XynB3 for saccharification.

Figure 9: EXP formation during an SSF reaction using an enzyme complex produced from the integrated *T. reesei* strain #229 in the presence of *T. reesei* Bxl1, or in the presence of certain other GH43 family β-xylosidase enzymes.

Figure 10: Reduced EXP formation observed from the addition of Fv43D under the recombinant *Zymomonas* SSF conditons.

Figure 11: Reduced EXP formation observed from the addition of Fv43D under the yeast SSF conditons.

Figure 12: Reduced EXP formation observed after the addition of Fv43D to Accellerase™ 1500+Multifect® Xylanase, to Accellerase™ 1500+XlnA, and to an

enzyme complex produced from the integrated *T. reesei* strain #229 under the recombinant *Zymomonas* SSF conditions.

Figure 13: Reduced EXP formation observed from the addition of Fv43D to purified cellulase enzymes and XynB3.

Figures 14A-14B: Figure 14A shows EXP reduction dose response from the addition of Fv43D to the enzyme composition or blend produced from the *T. reesei* integrated strain #229 + *T. reesei* Bxl1 under the recombinant *Zymomonas* SSF conditions. Figure 14B shows EXP reduction from the addition of Fo43A or Gz43A to the enzyme complex produced from the integrated *T. reesei* strain #229 + *T. reesei* Bxl1 under the recombinant *Zymomonas* SSF conditions.

Figure 15: Time course of xylose and EXP formation (expressed as RI area, which is proportional to mg/mL) from xylobiose (20 mg/mL) in 50 mM sodium citrate, pH 4.7 plus 0.9 M ethanol at 46 ºC in the presence of Multifect® Xylanase (560 µg/mL) and purified Fv43D (36 µg/mL) and Fv3A (54 µg/mL).

Figure 16: Time course of formation of xylose and EXP from xylobiose (20 mg/mL, left) or xylose oligomers (20 mg/mL, right), in 50 mM sodium citrate, pH 4.7 plus 0.9 M ethanol at 46 ºC in the presence of Multifect® Xylanase (560 µg/mL), Fv43D (36 µg/mL), or Fv3A (54 µg/mL). The results are expressed as ratios of the amount of EXP to the amount of xylose formed.

Figure 17: Plasmid map of pCR-Blunt II-TOPO, *bxl1* deletion, *hph-loxP*.

Figure 18: Plasmid map of TOPO Blunt/Pegl1-Fv43D.

Figures 19A-19B: Figure 19A: Fv43D nucleotide sequence (SEQ ID NO:1). Figure 19B: Fv43D amino acid sequence (SEQ ID NO:2). SEQ ID NO:2 is the sequence of the immature Fv43D. Fv43D has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:2 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 21 to 350 of SEQ ID NO:2. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk). The predicted conserved domain residues are in boldface type. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figures 20A-20B: Figure 20A: *T. reesei* Bxl1 nucleotide sequence (SEQ ID NO:3). Figure 20B: *T. reesei* Bxl1 amino acid sequence (SEQ ID NO:4). The signal sequence is underlined. The predicted conserved domain residues are in bold face type. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figures 21A-21B: Figure 21A: Fv3A nucleotide sequence (SEQ ID NO:5). Figure 21B: Fv3A amino acid sequence (SEQ ID NO:6). SEQ ID NO:6 is the sequence of the immature Fv3A. Fv3A has a predicted signal sequence corresponding to residues 1 to 23 of SEQ ID NO:6 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 24 to 766 of SEQ ID NO:6. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk). The predicted conserved domain residues are in boldface type. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figures 22A-22B: Figure 22A: Pf43A nucleotide sequence (SEQ ID NO:7). Figure 22B: Pf43A amino acid sequence (SEQ ID NO:8). SEQ ID NO:8 is the sequence of the immature Pf43A. Pf43A has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:8 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 21 to 445 of SEQ ID NO:8. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk). The predicted conserved domain residues are in boldface type, the predicted carbohydrate binding module ("CBM") residues are in uppercase type, and the predicted linker separating the CD and CBM is in italics. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figures 23A-23B: Figure 23A: Fv43E nucleotide sequence (SEQ ID NO:9). Figure 23B: Fv43E amino acid sequence (SEQ ID NO:10). SEQ ID NO:10 is the sequence of the immature Fv43E. Fv43E has a predicted signal sequence corresponding to residues 1 to 18 of SEQ ID NO:10 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 19 to 530 of SEQ ID NO:10. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk). The predicted conserved domain residues are in boldface type. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figures 24A-24B: Figure 24A: Fv43B nucleotide sequence (SEQ ID NO:11). Figure 24B: Fv43B amino acid sequence (SEQ ID NO:12). SEQ ID NO:12 is the sequence of the immature Fv43B. Fv43B has a predicted signal sequence corresponding to residues 1 to 16 of SEQ ID NO:12 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 17 to 574 of SEQ ID NO:12. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk). The predicted conserved domain

residues are in boldface type. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figures 25A-25B: Figure 25A: Af43A nucleotide sequence (SEQ ID NO:13). Figure 25B: Af43A amino acid sequence (SEQ ID NO:14). SEQ ID NO:14 is the sequence of the immature Af43A. Af43A does not have a predicted signal sequence, which can be derived using the SignalP algorithm (available at: http://www.cbs.dtu.dk). The predicted conserved domain residues are in boldface type. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figures 26A-26B: Figure 26A: Fv51A nucleotide sequence (SEQ ID NO:15). Figure 26B: Fv51A amino acid sequence (SEQ ID NO:16). SEQ ID NO:16 is the sequence of the immature Fv51A. Fv51A has a predicted signal sequence corresponding to residues 1 to 19 of SEQ ID NO:16 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 20 to 660 of SEQ ID NO:16. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk). The predicted L-α-arabinfuranosidase conserved domain residues are in boldface type. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figures 27A-27B: Figure 27A: *T. reesei* Xyn3 nucleotide sequence (SEQ ID NO:17). Figure 27B: *T. reesei* Xyn3 amino acid sequence (SEQ ID NO:18). SEQ ID NO:18 is the sequence of the immature *T. reesei* Xyn3. *T. reesei* Xyn3 has a predicted signal sequence corresponding to residues 1 to 16 of SEQ ID NO:18 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 17 to 347 of SEQ ID NO:18. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk). The predicted conserved domain residues are in bold face type. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figures 28A-28B: Figure 28A: XlnA nucleotide sequence (SEQ ID NO:19). Figure 28B: XlnA amino acid sequence (SEQ ID NO:20). SEQ ID NO:20 is the sequence of the immature XlnA protein. XlnA has a predicted signal sequence corresponding to residues 1 to 27 of SEQ ID NO:20 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 28 to 211 of SEQ ID NO:20. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk). SEQ ID NO:19 is the genomic sequence of XlnA; the initiation and termination codon residues are shown in bold face type in Figure 28A, and intron A of the XlnA gene is underlined in Figure 28A.

Figure 29: Figure 29 shows the "α" and "β" anomer configurations of glucose. Anomers are identified as "α" or "β" based on the relation between the stereochemistry of the exocyclic oxygen atom at the anomeric carbon and the oxygen attached to the configurational atom (defining the sugar as D or L), which is often the furthest chiral center in the ring. The α anomer is the one in which these two positions have the same configuration; they are the opposite in the β anomer. Thus the structure of α-D-glucose has the same stereochemistry at both C1 and C5 whereas β-D-glucose has opposite stereochemistry at C1 compared to C5.

Figures 30A-30B: Figure 30A: Gz43A nucleotide sequence (SEQ ID NO:21). Figure 30B: Gz43A amino acid sequence (SEQ ID NO:22). SEQ ID NO:22 is the sequence of the immature Gz43A. Gz43A has a predicted signal sequence corresponding to residues 1 to 18 of SEQ ID NO:22 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 19 to 340 of SEQ ID NO:22. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk). The predicted conserved domain residues are in boldface type. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figures 31A-31B: Figure 31A: Fo43A nucleotide sequence (SEQ ID NO:23). Figure 31B: Fo43A amino acid sequence (SEQ ID NO:24). SEQ ID NO:24 is the sequence of the immature Fo43A. Fo43A has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:24 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 21 to 348 of SEQ ID NO:24. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk). The predicted conserved domain residues are in boldface type. Domain predictions were made based on the Pfam, SMART, or NCBI databases.

Figure 32-1 to 32-2: Alignment of GH43 family hydrolases. Amino acid residues that are highly conserved among members of the family are shown in bold and underline type.

Figure 33: XynB3 amino acid sequence (SEQ ID NO:25).

Figure 34: *T. reesei* Bgl1 amino acid sequence (SEQ ID NO:45). The signal sequence is underlined. The predicted conserved domain residues are in bold face type. The coding sequence is described in Barnett *et al.*, 1991, Bio-Technology 9(6):562-567.

Figures 35A-35B: Figure 35A: XlnD nucleotide sequence (SEQ ID NO:39). Figure 35B: XlnD amino acid sequence (SEQ ID NO:40). SEQ ID NO:40 is the

sequence of the immature XlnD. XlnD has a predicted signal sequence corresponding to residues 1 to 17 of SEQ ID NO:40 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 18 to 804 of SEQ ID NO:40. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk).

Figures 36A-36B: Figure 36A: Fv30A nucleotide sequence (SEQ ID NO:41). Figure 36B: Fv30A amino acid sequence (SEQ ID NO:42). SEQ ID NO:42 is the sequence of the immature Fv30A. Fv30A has a predicted signal sequence corresponding to residues 1 to 19 of SEQ ID NO:42(underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 20 to 537 of SEQ ID NO:42. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk).

Figures 37A-37B: Figure 37A: Fv30B nucleotide sequence (SEQ ID NO:43). Figure 37B: Fv30B amino acid sequence (SEQ ID NO:44). SEQ ID NO:44 is the sequence of the immature Fv30B. Fv30B has a predicted signal sequence corresponding to residues 1 to 24 of SEQ ID NO:44 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 25 to 485 of SEQ ID NO:44. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk).

Figures 38A-38B: Figure 38A: Fv39A nucleotide sequence (SEQ ID NO:45). Figure 38B: Fv39A amino acid sequence (SEQ ID NO:46). SEQ ID NO:46 is the sequence of the immature Fv39A. Fv39A has a predicted signal sequence corresponding to residues 1 to 19 of SEQ ID NO:46 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 20 to 439 of SEQ ID NO:46. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk).

Figures 39A-39B: Figure 39A: Fv39B nucleotide sequence (SEQ ID NO:47). Figure 39B: Fv39B amino acid sequence (SEQ ID NO:48). SEQ ID NO:48 is the sequence of the immature Fv39B. Fv39B has a predicted signal sequence corresponding to residues 1 to 18 of SEQ ID NO:48 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 19 to 456 of SEQ ID NO:48. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk).

Figures 40A-40B: Figure 40A: XynB nucleotide sequence (SEQ ID NO:49). Figure 40B: XynB amino acid sequence (SEQ ID NO:50). XynB does not have a

predicted signal sequence from the SignalP algorithm (available at: http://www.cbs.dtu.dk).

Figures 41A-41B: Figure 41A: XylA nucleotide sequence (SEQ ID NO:51). Figure 41B: XylA amino acid sequence (SEQ ID NO:52). XylA does not have a predicted signal sequence from the SignalP algorithm (available at: http://www.cbs.dtu.dk), but has a signal sequence predicted from the Uniprot algorithm (available at: http://www.uniprot.org/uniprot) that corresponds to residues 1 to 18 of SEQ ID NO:52 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 19-705 of SEQ ID NO:52.

Figures 42A-42B: Figure 42A: Xyl1 nucleotide sequence (SEQ ID NO:53). Figure 42B: Xyl1 amino acid sequence (SEQ ID NO:54). SEQ ID NO:54 is the sequence of the immature Xyl1. Xyl1 has a predicted signal sequence corresponding to residues 1 to 21 of SEQ ID NO:54 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 22 to 500 of SEQ ID NO:54. Signal sequence predictions were made with the SignalP algorithm (available at: http://www.cbs.dtu.dk).

Figures 43A-43B: Figure 43A: EXP and ethanol concentrations measured on Day 1 from an SSF reaction employing the *bxl1*⁻ *T. reesei* strain #229, where 0.5 or 1.5 mg/g xylan of purifieid *T. reesei* Bxl1 was added to the SSF reaction or where 1 mg/g xylan of purified Fv43D was added; Figure 43B: EXP and ethanol concentrations measured on Day 3 from an SSF reaction employing the *bxl1*⁻ *T. reesei* #229, where 0.5 or 1.0 mg/g xylan or purified *T. reesei* Bxl1 was added to the SSF reaction or where 1 mg/g xylan of purified Fv43D was added. Conditions of the SSF reaction(s) are described below in Example 6.

6. DETAILED DESCRIPTION

The meanings of abbreviations used herein are listed below: "min" means minute, "mins" means minutes; "hr" means hour, "hrs" means hours, "d" means day(s), "µL" means microliter(s), "mL" means milliliter(s), "L" means liter(s), "nm" means to nanometer(s), "mm" means millimeter(s), "cm" means centimeter(s), "µm" means micrometer(s), "mM" means millimolar(s), "M" means molar(s), "mmoL" means millimole(s), "µmole" means micromole(s), "g" means gram(s), ""µg" means microgram(s), "mg" means to milligram(s), "kg" means kilogram(s), , "RPM" or "rpm" means revolutions per minute, "vol.%" means volume %, "wt.%" means weight %, and "RPS" means revolutions per second.

6.1. Common Definitions

Unless otherwise noted, all U.S. Patents and U.S. Patent Applications cited to herein are incorporated by reference in their entirety. Moreover, when an amount, concentration, or other value or parameter is given as a range, a preferred range, or a list of upper preferable value, or lower preferable values, it should be understood as specifically disclosing all ranges or numbers along the continuum formed in those ranges. When a range of numerical values is recited herein, unless otherwise noted, the range is intended to encompass the endpoints of that range, and all intergers and fractions within the range. It is not intended that the scope of the invention be limited to the specific values recited when defining a range.

As used herein, the articles "a", "an", and "the" preceding an element or component of the invention are intended to be nonrestrictive regarding the number of instances (e.g., occurrences) of the element or component. Thus "a", "an", and "the" should be read to include one or at least one, and the singular word form of the element or component also includes the plural unless the number is obviously meant to be singular.

As used herein, the term "comprising" means the presence of the stated features, integers, steps, or components as referred to in the claims, but that it does not preclude the presence or addition of one or more other features, integers, steps, components or groups thereof. The term "comprising" is intended to include embodiments encompassed by the term "consisting essentially of" and "consisting of." Similarly, the term "consisting essentially of" is intended to include embodiments encompassed by the term "consisting of."

As used herein, the term "about" modifying the quanity of an ingredient or reactant, or the quantity of a parameter of the invention, refers to variation in the numerical quantity that can occur, for example, through typical measuring and liquid handling procedures used for making concentrates or solutions in the real world, through inadvertent errors in these procedures, through differences in the manufacture, source, or purity of the ingredients employed to make the compositions or carry out the methods; and the like. The term "about" also encompasses amounts that differ due to different equilibrium conditions for a composition resulting from a particular initial mixture. Whether or not modified by the term "about" the claims include equivalents to the quanities they recite.

The term "simultaneous saccharification and fermentation" or "SSF" refers to a process or reaction configuration wherein biomass is saccharified and the fermentable

sugars produced from the saccharification are used by an enzyme and/or by a fermenting microorganism to produce a product all at the same time, typically in the same reaction vessel.

The term "hybrid saccharification and fermentation" or "HSF" refers to a process or reaction configuration wherein biomass is saccharified to a limited extent (incomplete or partial saccharification), followed by continued saccharification and fermentation occurring simultaneously.

The terms "separate saccharification and fermentation," "separate hydrolysis and fermentation," and "SHF" are used interchangeably herein. They refer to a process or reaction configuration wherein biomass is saccharified or hydrolyzed to substantial completion (*e.g.,* about 60% or more complete, about 70% or more complete, about 80% or more complete, about 90% or more complete, or about 95% or more complete) or to completion (*e.g.,* about 99% or more complete, or about 100% complete, such that all fermentable sugars that would be released from a given saccharification reaction are released), followed by a separate and distinct fermentation step, wherein the fermentation sugars produced by the saccharification or hydrolysis step is fermented to produce a fermentation product.

The term "fermentable sugar" refers to oligosaccharides and monosaccharides that can be used as a carbon source by a microorganism in a fermentation process.

The term "partial saccharifcation" refers to limited saccharification of biomass where the fermentable sugars released are less than the total of fermentable sugars that would be released if saccharification is run to completion.

The term "cellulosic" refers to a composition comprising cellulose and additional components, including, for example, hemicellulose.

The term "saccharification" refers to the production of fermentable sugars from polysaccharides or polysaccharide-containing materials.

The term "biomass" refers to any cellulosic or lignocellulosic material and includes materials comprising cellulose, and optionally further comprising hemicellulose, lignin, starch, oligosaccharides, and/or monosaccharides. Biomass can also comprise additional components, such as proteins and/or lipids. Biomass can be derived from a single source, or biomass can comprise a mixture derived from more than one source. For example, biomass can comprise a mixture of corn cobs and corn stover, or a mixture of grass and leaves. Biomass incudes, but is not limited to, bioenergy crops, agricultural residues, municipal solid waste, industrial solid waste, sludge from paper manufacture, yard waste, wood and forestry waste. Examples of biomass include, without limitation,

corn cobs, crop residues such as corn husks, corn stover, grasses, wheat, wheat straw, barley straw, hay, rice straw, switchgrass, wasted paper, sugar cane bagasse, sorghum, giant reed, elephant grass, miscanthus, Japanese cedar, components obtained from milling of grains, tress, branches, roots, leaves, wood chips, sawdust, shrubs and bushes, vegetables, fruits, flowers and animal manure.

The term "saccharification enzyme" refers to an enzyme that can catalyze conversion of a component of biomass to fermentable sugars. It is often the case that the enzyme is more effective at producing fermentable sugars when the biomass is pretreated.

6.2. Detailed Description

Production of a substance or fermentation product from cellulosic material typically involves three major steps. These three steps are (1) pretreatment or pre-hydrolysis, (2) enzymatic hydrolysis or saccharification, and (3) fermentation, after which the substance or fermentation product can be recovered. Exemplified below is a process for producing ethanol, but it will be understood that similar processes can be used to produce other substances.

Pretreatment. In the pretreatment or pre-hydrolysis step, the cellulosic material (including, for example, a lignocellulosic material) is heated to break down the lignin and carbohydrate structure, to solubilize most of the hemicellulose, and to make the cellulose fraction accessible to cellulolytic enzymes. The heating step is performed either directly using steam or in a slurry or mixture where a catalyst may optionally be added to the material to accelerate the reactions. Suitable catalysts include, for example, strong acids, such as sulfuric acid and SO_2, or strong bases, such as sodium hydroxide. The pretreatment step facilitates the penetration of the enzymes and microorganisms. Cellulosic biomass may also be subject to a hydrothermal steam explosion pre-treatment (*see, e.g.*, U.S. Patent Publication No. 2002/0164730).

Saccharification. In the enzymatic hydrolysis step, also known as the saccharification step, enzymes as described herein are added to the pretreated material to convert the cellulose fraction to glucose and/or other sugars. The saccharification step is generally performed in stirred-tank reactors or fermentors under controlled pH, temperature, and mixing conditions. A saccharification step may, in certain cases, last up to 200 hrs. Saccharification can be carried out at temperatures from about 30 °C to about 65 °C, in particular about 50 °C, and at a pH of between about 4 and about 5, in particular at about pH 4.5. To produce glucose that can be metabolized by a fermenting microorganism such as a fungus (e.g., a yeast or a filamentous fungus) or a bacterium

(e.g., a *Zymomonas mobillis* or a *Clostridium thermocellum*) the enzymatic hydrolysis step is typically performed in the presence of a β-glucosidase.

Fermentation. In the fermentation step, sugars, released from the cellulosic material as a result of the pretreatment and enzymatic hydrolysis steps, are fermented to ethanol (or other substances) by a fermenting organism, such as a fungus (e.g., a yeast or a filamentous fungus) or a bacterium (e.g., a *Zymomonas mobillis* or a *Clostridium thermocellum*).

SSF. The present disclosure provides methods and compositions for improving the yield of reactions in which the fermentation step is carried out, rather than in a distinct or separate step following the enzymatic hydrolysis step, simultaneously with the enzymatic hydrolysis step in the same vessel, preferably under controlled pH, temperature, and mixing conditions. In certain aspects, the saccharification and fermentation are performed simultaneously in the same vessel, and as such is a simultaneous saccharification and fermentation, or "SSF." This process, as described herein, encompasses, also processes that are carried out using a "hybrid saccharification and fermentation" or "HSF" configuration. In certain aspects, an SSF reaction is initiated (*e.g.*, by the addition of fermenting microorganism to a saccharification reaction, or by instituting a set of conditions to favor fermentation) when no more than 30%, no more than 25%, no more than 20%, no more than 15%, no more than 10%, or no more than 5% of the biomass is saccharified. As used herein, the term "SSF" also encompasses the co-fermentation of multiple sugars (Sheehan and Himmel, 1999, Enzymes, energy and the environment: A strategic perspective on the U.S. Department of Energy's research and development activities for bioethanol, Biotechnol. Prog. 15: 817-827).

6.3 Enzymatic Hydrolysis

The cell walls of higher plants are comprised of a variety of carbohydrate polymer (CP) components. These CP components interact through covalent and non-covalent means, providing the structural integrity plants require to form rigid cell walls and to resist turgor pressure. The major CP found in plants is cellulose, which forms the structural backbone of the plant cell walls. 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 somewhat irregular and contain regions of varying crystallinity. The degree of crystallinity of cellulose fibrils depends on how tightly ordered the hydrogen bonding is between any two component cellulose chains. Areas with less-ordered

bonding, and therefore more accessible glucose chains, are referred to as amorphous regions.

The general model for converting or depolymerizing cellulose into glucose involves three enzymatic activities. Endoglucanases cleave cellulose chains internally to generate shorter chains and increase the number of accessible ends, which are then acted upon by exoglucanases. Exoglucanases are specific for either the reducing ends or the non-reducing ends of the shorter chains, and are capable of liberating cellobiose, the dimer of glucose. Examples of exoglucanases include, without limitation, various cellobiohydrolases. The accumulating cellobiose is then cleaved to form glucose by cellobiases. Examples of cellobiases include, without limitation, various β-1,4-glucosidases.

Hemicellulose contains a set of different sugar monomers from those of cellulose, which contains anhydro-glucose. For instance, aside from glucose, sugar monomers in hemicellulose can include xylose, mannose, galactose, rhamnose, and arabinose. Hemicelluloses contain mostly D-pentose sugars and occasionally small amounts of L-sugars as well. Xylose is the sugar monomer present in the largest amount, but mannuronic acid and galacturonic acid also tend to be present. Hemicelluloses include, for example, xylan, glucuronoxylan, arabinoxylan, glucomannan, and xyloglucan.

Enzymes and multi-enzyme compositions of the present disclosure are useful for saccharification of hemicellulose materials, *e.g.*, xylan, arabinoxylan and xylan- or arabinoxylan-containing substrates. Arabinoxylan is a polysaccharide composed of xylose and arabinose, wherein L-α -arabinofuranose residues are attached as branch-points to a β-(1,4)-linked xylose polymeric backbone.

Due to the complexity of most biomass sources, which can contain cellulose, hemicellulose, pectin, lignin, protein, and ash, among other components, in certain aspects the enzyme blends of the disclosure can contain enzymes with a range of substrate specificities that work together to degrade biomass into fermentable sugars in an efficient manner. One example of a multi-enzyme complex for lignocellulose saccharification comprises a mixture of cellobiohydrolase(s), xylanase(s), endoglucanase(s), β-glucosidase(s), β-xylosidase(s), and, optionally, various accessory proteins.

Accordingly, the present disclosure contemplates the use of one or more enzymes that are capable, individually or collectively, of producing a carbohydrate that can be used as an energy-source by the fermenting organism(s) in an SSF reaction for producing a fermentation product, such as ethanol.

In certain aspects, multi-enzyme compositions are used in an SSF reaction for hydrolysis of carbohydrates or carbohydrate-containing biomass substrates to produce sugars that are fermented in the same reaction by a fermenting microorganism. The multi-enzyme compositions (including products of manufacture, enzyme ensembles, or "blends") comprise a mixture (or "blend") of enzymes that, in certain aspects, is non-naturally occurring. As used herein, the term "blend" refers to:

(1) a composition made by combining component enzymes, whether in the form of a fermentation broth or in the form of partially or completely isolated or purified polypeptides;

(2) a composition produced by an organism modified to express one or more component enzymes; optionally, the organism can be modified to delete one or more genes or inactivate one or more gene products, wherein the genes encode proteins that affect xylan hydrolysis, hemicellulose hydrolysis and/or cellulose hydrolysis;

(3) a composition made by combining component enzymes simultaneously, separately, or sequentially during an SSF reaction; and

(4) an enzyme mixture produced *in situ*, *e.g.*, during an SSF reaction;

(5) a combination of any or all of (1)-(4) above.

It is also to be understood that 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 multi-enzyme composition. The disclosure is not restricted or limited to the specific exemplary combinations listed or exemplified herein.

In the methods of the present disclosure, any of the enzyme(s) described herein can be added prior to or during the SSF reaction, including during or after the propagation of the fermenting microorganism(s). The enzymes can be added individually, as an enzyme blend, or as a fermentation broth, *etc.*

The enzymes referenced herein can be derived or obtained from any suitable origins, including, for example, from bacterial, fungal, yeast, or mammalian origins. The term "obtained" is meant that the enzyme can be isolated from an organism, which naturally produces the enzyme as a native enzyme, or that the enzyme can be produced recombinantly in a host organism, wherein the recombinantly produced enzyme is either native or foreign to the host organism, has a modified amino acid sequence, *e.g.*, having one or more amino acids, which are deleted, inserted and/or substituted, or is an enzyme produced by nucleic acid shuffling processes known in the art. For example, the

recombinantly produced enzyme can be one that is a mutant and/or a fragment of a native amino acid sequence. By “a native enzyme” it is meant to encompass the product of a gene in its natural location in the genome of an organism, and also to encompass natural variants; by a “foreign enzyme” it is meant to encompass the product of a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer, or gene inserted into a non-native organism, or chimeric genes, which includes enzymes that are obtained recombinantly, such as by site-directed mutagenesis or shuffling.

The enzymes can, in certain aspects, be purified. The term “purified," as used herein to modify substances like enzymes, proteins, polypeptides, polynucleotides, and other components, refers to enzymes free or substantially free from other components of the organisms from which they are derived. In certain aspects, the term "purified" also encompasses situations wherein the enzymes are free or substantially free from components of the native organisms from which they are obtained. Enzymes can be deemed “purified,” but remain associated or in the presence of minor amounts of other proteins. The term "other proteins," as used herein, refers, in particular, to other enzymes. The term "purified" as used herein also refers to the removal of other components, particularly the removal of other proteins and more particularly other enzymes present in the originating cells of the enzymes of the disclosure. Accordingly, an enzyme can be, for example, a "substantially pure polypeptide," which is substantially free from other components. The organism in which a given enzyme is produced can be, for example, a host organism suitable for recombinantly produced enzymes. For example, a substantially pure polypeptide can refer to a polypeptide present at a level of 50 wt.% or more, 60 wt.% or more, 70 wt.% or more, 80 wt.% or more, 90 wt.% or more, 95 wt.% or more, 98% or more, or 99% or more in a mixture to which it is a part. A polypeptide substantially free of other components is one that is in a mixture that contains less than 40 wt.%, less than 30 wt.%, less than 20 wt.%, less than 10 wt.%, less than 5 wt.%, less than 2 wt.%, or less than 1 wt.% of other components.

6.3.1. Cellulases

Enzyme blends of the disclosure can comprise one or more cellulases. Cellulases are enzymes that hydrolyze cellulose (β-1,4-glucan or β-D-glucosidic linkages) to form glucose, cellobiose, cellooligosaccharides, and the like. Cellulases are 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 Biotechnol.

5(9):255-261, and Schulein, 1988, Methods of Enzymology 160:234-242). Endoglucanases act mainly on the amorphous parts of the cellulose fiber, whereas cellobiohydrolases are capable of degrading crystalline cellulose.

Cellulases for use in accordance with the methods and compositions of the disclosure can be obtained 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*, *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 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.*

In specific embodiments, a cellulase for use in the composition of the disclosure is capable of achieving at least 0.1, at least 0.2, at least 0.3, at least 0.4, or at least 0.5 fraction product as determined by a calcofluor assay as described in the following subsection. In some embodiments, a cellulase for use in the composition of the disclosure is a whole cellulase and/or is capable of achieving at least 0.1, at least 0.2, at least 0.3, at least 0.4, or at least 0.5 fraction product as determined by a calcofluor assay as described in the following subsection. In some embodiments, a cellulase for use in the composition of the disclosure is a whole cellulase and/or is capable of achieving about 0.1 to about 0.5, or about 0.1 to about 0.4, or about 0.2 to about 0.4, or about 0.3 to about 0.4, or about 0.2 to about 0.5, or about 0.3 to about 0.5 fraction product as determined by a calcofluor assay.

6.3.1.1. Cellulase Activity Assay Using Calcofluor White

Phosphoric acid swollen cellulose (PASC) is prepared from Avicel PH-101 using an adapted protocol of Walseth, 1971, TAPPI 35:228, and of Wood, 1971, Biochem. J. 121:353-362. In short, in an exemplary method, Avicel is solubilized in concentrated

phosphoric acid then precipitated using cold deionized water. After the cellulose is collected and washed with water to achieve neutral pH, it is diluted to 1% solids in 50 mM sodium acetate buffer, at pH 5.0.

All enzyme dilutions are made with a 50 mM sodium acetate buffer, pH 5.0. GC220 Cellulase (Danisco US Inc., Genencor) is diluted to 2.5, 5, 10, and 15 mg protein/g PASC, to produce a linear calibration curve. Samples to be tested are diluted to fall within the range of the calibration curve, *i.e.*, to obtain a response of 0.1 to 0.4 fraction product. One hundred and fifty (150) µL of cold 1% PASC is added to each 20 µL of enzyme solution in suitable vessles, for example, 96-well microtiter plates. The plates are covered and incubated for 2 hrs at 50 °C, spun at 200 rpm, in an incubator/shaker. The reactions are then quenched using 100 µL of 50 µg/mL Calcofluor in 100 mM Glycine, pH 10. Fluorescence is read on a fluorescence microplate reader 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 sample - FI buffer with cellobiose)/(FI zero enzyme – FI buffer with cellobiose), wherein "FP" is fraction product and "FI" is fluorescence units.

6.3.1.2. β-Glucosidase

The enzyme blends of the disclosure optionally comprise one or more β-glucosidases. The term "β-glucosidase" as used herein refers to a β-D-glucoside glucohydrolase classified in or under EC 3.2.1.21, and/or to an enzyme that is a member of certain glycosyl hydrolase ("GH") families, including, without limitation, GH families 1, 3, 9 or 48. In certain aspects, the term refers to an enzyme that is capable of catalyzing the hydrolysis of cellobiose to release β-D-glucose.

β-glucosidases can be obtained from any suitable microorganisms. They can be obtained or produced by recombinant means, or can be obtained from commercial sources. Suitable β-glucosidases can, for example, be obtained from microorganisms such as bacteria and fungi. For example, a suitable β-glucosidase can be obtained from a filamentous fungus.

In certain aspects, a suitable β-glucosidase can be obtained from *Aspergillus aculeatus* (Kawaguchi *et al.*, 1996, Gene 173: 287-288), *Aspergillus kawachi* (Iwashita *et al.*, 1999, Appl. Environ. Microbiol. 65: 5546-5553), *Aspergillus oryzae* (PCT patent application publication WO 2002/095014), *Cellulomonas biazotea* (Wong *et al.*, 1998, Gene 207:79-86), *Penicillium funiculosum* (PCT patent application publication WO 200478919), *Saccharomycopsis fibuligera* (Machida *et al.*, 1988, Appl. Environ. Microbiol. 54: 3147-3155), *Schizosaccharomyces pombe* (Wood *et al.*, 2002, Nature

415: 871-880), or *Trichoderma reesei*. For example, suitable β-glucosidases from *Trichoderma reesei* can include β-glucosidase 1 (U.S. Patent No. 6,022,725), *Trichoderma reesei* β-glucosidase 3 (U.S. Patent No. 6,982,159), *Trichoderma reesei* β-glucosidase 4 (U.S. Patent No. 7,045,332), *Trichoderma reesei* β-glucosidase 5 (US Patent No. 7,005,289), *Trichoderma reesei* β-glucosidase 6 (U.S. Patent Application Publication 20060258554), or *Trichoderma reesei* β-glucosidase 7 (U.S. Patent Application Publication 20060258554).

In some embodiments, suitable β-glucosidases can be produced by expressing genes encoding β-glucosidases. For example, a suitable β-glucosidase can be secreted into the extracellular space *e.g.*, by certain Gram-positive organism, (such as *Bacillus* or *Actinomycetes*), or by a eukaryotic host (*e.g.*, *Trichoderma*, *Aspergillus*, *Saccharomyces*, or *Pichia*).

Suitable β-glucosidases can also be obtained from commercial sources. Examples of commercial β-glucosidase preparation suitable for use in the methods, compositions and other embodiments of the present disclosure include, without limitation, *Trichoderma reesei* β-glucosidase in Accellerase™ BG (Danisco US Inc., Genencor); NOVOZYM™ 188 (a β-glucosidase from *Aspergillus niger*); *Agrobacterium sp.* β-glucosidase, and *Thermatoga maritima* β-glucosidase available from Megazyme (Megazyme International Ireland Ltd., Bray Business Park, Bray,Co. Wicklow, Ireland.).

In certain aspects, a suitable β-glucosidase can be a component of a whole cellulase, as described in Section 6.3.1.5 below.

β-glucosidase activity can be determined by means that are known in the art. For example, the assay described by Chen *et al.*, 1992, in Biochimica et Biophysica Acta 121:54-60 can be used. In that assay, one unit pNPG denotes 1 μmoL of Nitrophenol liberated from para-nitrophenyl-B-D-glucopyranoside in 10 mins at 50 °C (or 122 °F) and pH 4.8.

6.3.1.3. Endoglucanases

The enzyme blends of the disclosure optionally comprise one or more endoglucanase. The term "endoglucanase" refers to any polypeptides classified in EC 3.2.1.4..

In some aspects, a *Trichoderma reesei* EG1 (Penttila *et al.*, 1986, Gene 63:103-112) and/or *T. reesei* EGII (Saloheimo *et al.*, 1988, Gene 63:11-21) are used in the methods and compositions of the present disclosure. In other aspects, the endoglucanase can be a *T. reesei* endoglucanase VI (*see, e.g.*, U.S. Patent No.

7,351,568), endoglucanase VII (*see, e.g.*, U.S. Patent No. 7,449,319), or endoglucanase VIII (*see, e.g.*, U.S. Patent No. 7,049,125).

In specific embodiments, a suitable endoglucanase can be a *Thielavia terrestris* thermostable endoglucanase (Kvesitadaze *et al.*, 1995, Appl. Biochem. Biotechnol. 50:137-143); *Trichoderma reesei* EGIII (Okada *et al.*, 1988, Appl. Environ. Microbiol. 64:555-563), EGIV (Saloheimo *et al.*, 1997, Eur. J. Biochem. 249:584-591), EG5 (Saloheimo *et al.*, 1994, Mol. Microbiol. 13:219-228), EGVI (U.S. Patent Application Publication No. 20070213249), or EGVII (U.S. Patent Application Publication No. 20090170181); *Acidothermus cellulolyticus* EI endoglucanase (U.S. Patent No. 5,536,655); *Humicola insolens* endoglucanase V (EGV) (Protein Data Bank entry 4ENG); *Staphylotrichum coccosporum* endoglucanase (U.S. Patent Application Publication No. 20070111278); *Aspergillus aculeatus* endoglucanase F1-CMC (Ooi *et al.*, 1990, Nucleic Acid Res. 18:5884); Aspergillus *kawachii* IFO 4308 endoglucanase CMCase-1 (Sakamoto *et al.*,1995, Curr. Genet. 27:435-439); or *Erwinia carotovara* (Saarilahti *et al.*, 1990, Gene 90:9-14); *Acremonium thermophilum* ALKO4245 endoglucanase (U.S. Patent Publication No. 20070148732).

Suitable endoglucanases for use in the methods and compositions of the present disclosure can also be those described in, *e.g.*, PCT patent application publications WO 91/17243, WO 91/17244, WO 91/10732, or U.S. Patent No. 6,001,639.

6.3.1.4. Cellobiohydrolases

The term "cellobiohydrolase" as used herein refers to any cellobiohydrolases that are classified in EC 3.2.1.91. The methods and compositions of the present disclosure can suitably comprise one or more cellobiohydrolases ("CBH").

In some aspects, a *Trichoderma reesei* CBHI (Shoemaker *et al.*, 1983, Bio/Technology 1:691-696) and/or CBHII (Teeri *et al.*, 1983, Bio/Technolgy 1:696-699) can be used in the methods and compositions of the present disclosure.

In some aspects, a suitable CBH can be an *Agaricus bisporus* CBH1 (Swiss Prot Accession no. Q92400); *Aspergillus aculeatus* CBH1 (Swiss Prot Accesion No. O59843); *Aspergillus nidulans* CBHA (GenBank Accession No. AF420019); *Aspergillus nidulans* CBHB (GenBank Accession No. AF420020); *Aspergillus niger* CBHA (GenBank Accession No. AF156268); *Aspergillus niger* CBHB (GenBank Accession No. AF156269); *Claviceps purpurea* CBH1 (Swiss Prot Accession No. O00082); *Cochliobolus carbonarum* CBH1 (Swiss Prot Accession No. Q00328); *Cryphonectria parasitica* CBH1 (Swiss Prot Accession No. Q00548); *Fusarium oxysporum* CBH1 (Cel7A) (Swiss Prot Accession No. P46238); *Humicola grisea* cbh1.2 (GenBank

Accession No. U50594); *Humicola grisea* var. *thermoidea* CBH1 (GenBank Accession No. D63515); *Humicola grisea* var. *thermoidea* CBHI.2 (GenBank Accession No. AF123441); *Humicola grisea* var. *thermoidea* exo1 (GenBank Accession No. AB003105); *Melanocarpus albomyces* Cel7B (GenBank Accession No. AJ515705), *Neurospora crassa* CBHI (GenBank Accession No. X77778); *Penicillium funiculosum* CBHI (Cel7A) (U.S. Patent Publication No. 20070148730); *Penicillium janthinellum* CBHI (GenBank Accession No. S56178); *Phanerochaete chrysosporium* CBH (GenBank Accession No. M22220); *Phanerochaete chrysosporium* CBHI-2 (Cel7D) (GenBank Accession No. L22656); *Talaromyces emersonii* Cbh1A (GenBank Accession No. AF439935); *Trichoderma viride* CBH1 (GenBank Accession No. X53931), or *Volvariella volvacea* V14 Cbh1 (GenBank Accession No. AF156693).

6.3.1.5. Whole Cellulases

In certain aspects, an enzyme blend of the disclosure comprises a whole cellulase. As used herein, a "whole cellulase" refers to both naturally occurring and non-naturally occurring cellulase-containing compositions comprising: (1) an endoglucanase, which cleaves internal β-1,4 linkages of a cellulose, resulting in shorter glucooligosaccharides, (2) a cellobiohydrolase, which acts in an "exo" manner to release cellobiose units from the shorter glucooligosaccharides; examples of cellobiose units include β-1,4 glucose-glucose disaccharide, and (3) a β-glucosidase, which catalyzes the release of glucose monomers from short cellooligosaccharides or cellobioses, which are glucose dimmers. .

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 ratios and levels 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 ratios 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. 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 filamentous fungal whole cellulase. For example, the whole cellulase preparation can be from an *Acremonium, Aspergillus, Emericella, Fusarium, Humicola, Mucor, Myceliophthora, Neurospora, Penicillium, Scytalidium, Thielavia, Tolypocladium,* or *Trichoderma* species. The whole cellulase preparation is, for example, an *Aspergillus aculeatus, Aspergillus awamori, Aspergillus foetidus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger,* or *Aspergillus oryzae* whole cellulase. Moreover, the whole cellulase preparation can be a *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*, or *Fusarium venenatum* whole cellulase preparation. The whole cellulase preparation can also be a *Chrysosporium lucknowense, Humicola insolens, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Penicillium purpurogenum, Penicillium funiculosum, Scytalidium thermophilum*, or *Thielavia terrestris* whole cellulase preparation. Moreover, the whole cellulase preparation can be a *Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei* (*e.g.,* RL-P37 (Sheir-Neiss G *et al.* Appl. Microbiol. Biotechnol. 1984, 20, pp.46-53), QM9414 (ATCC No. 26921), NRRL 15709, ATCC 13631, 56764, 56466, 56767), or a *Trichoderma viride* (*e.g.,* ATCC 32098 and 32086) whole cellulase preparation.

The whole cellulase preparation can, in particular, suitably be a *Trichoderma reesei* RutC30 whole cellulase preparation, which is available from the American Type Culture Collection as *Trichoderma reesei* ATCC 56765. For example, the whole cellulase preparation can also suitably be a whole cellulase of *Penicillium funiculosum*, which is available from the American Type Culture Collection as *Penicillium funiculosum* ATCC Number: 10446.

The whole cellulase preparation can also be obtained from commercial sources. 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).

Suitable whole cellulase preparations can be made using any microorganism cultivation methods known in the art, 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 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 and variations known in the art. Suitable culture media, temperature ranges and other conditions for growth and cellulase production are known in the art. As a non-limiting example, a typical temperature range for the production of cellulases by *Trichoderma reesei* is 24 °C to 28 °C

The whole cellulase preparation can be used as it is produced by fermentation with no or minimal recovery and/or purification. 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, for example, 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 methods known in the art.

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).

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, for example, comprise a whole cellulase preparation and a β-glucosidase. For instance, the β-glucosidase-enriched whole cellulase composition can suitably comprise at least 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.

In certain aspects, a suitable whole cellulase can be obtained from a microorganism that is or has been genetically engineered to reduce or eliminate retaining β-xylosidase activity. In other aspects, a suitable whole cellulase can be obtained from a microorganism that is or has been genetically engineered to increase inverting β-xylosidase activity. In yet further aspects, a suitable whole cellulase can be obtained from a microorganism that is or has been genetically engineered to not only have reduced or eliminated retaining β-xylosidase activity, but also have increased inverting β-xylosidase activity. For example, a whole cellulase can suitably be obtained from *Trichoderma reesei* that has been engineered such that the native *bxl1* gene is deleted. In another example, a whole cellulase can suitably be obtained from *Trichoderma reesei* that has been engineered to recombinantly express an enzyme with inverting β-xylosidase activity. In yet another example, a whole cellulase can suitably be obtained from *Trichoderma reesei* that has been engineered such that its native *bxl1* gene is deleted and that it recombinantly expresses an enzyme with inverting β-xylosidase activity. Examples of enzymes with inverting β-xylosidase activity include, without limitation, Fv43D and others described herein in Section 6.4.

β-xylosidase activity can be determined by measuring the level of hydrolysis of an artificial substrate *p*-nitrophenyl-β-xylopyranoside. The hydrolysis reaction can be followed using ^{1}H-NMR analysis during the course of the reaction. The anomeric proton of the residue contributing the reducing end of a glycosidic bond has a distinct chemical shift depending upon its axial or equatorial orientation as does the anomeric proton of the newly formed reducing sugar after hydrolysis. Mutarotation of the newly formed

reducing sugar anomeric proton to the equilibrium mixture of axial and equatorial forms is slower in comparison to the hydrolysis reaction. Thus the ^{1}H-NMR determination of the orientation of the first formed reducing end anomeric proton in comparison to the form present in the substrate is an assay for mechanisms that retain configuration versus those that invert configuration. The experimental methods are described in, *e.g.*, Pauly *et al.*, 1999, Glycobiology 9:93-100.

Alternatively, the level of hydrolysis can be determined by distinguishing transglycosylase activity of the retaining enzymes, which is absent in inverting enzymes. An example of such an assay is shown in Figure 16. Xylobiose or xylose oligomers, in the presence of a retaining enzyme (*e.g.*, Multifect® Xylanase or Fv3A), show a rapid rise in EXP/xylose to 7-8 times the equilibrium ratio in the presence of EtOH after which EXP/xylose falls toward the equilibrium ratio. In the case of the inverting enzymes (*e.g.*, Fv43D), under the same conditions, EXP/xylose increases monotonically toward the equilibrium ratio.

6.3.2. Hemicellulases

A wide variety of fungi and bacteria are capable of enzymatically hydrolyzing hemicelluloses. Similar to cellulose degradation, hemicellulose hydrolysis involves coordinated actions of a number of enzymes. Hemicellulases are often grouped into three general categories: the endo-acting enzymes that attack internal bonds within polysaccharide chains, the exo-acting enzymes that act processively from either the reducing or the nonreducing end of the polysaccharide chain, and the accessory enzymes, acetylesterases, and/or esterases that hydrolyze lignin glycoside bonds. Examples of esterases can include coumaric acid esterase and ferulic acid esterase (Wong *et al.*, 1988, Microbiol. Rev. 52:305-317; Tenkanen and Poutanen, 1992, Significance of esterases in the degradation of xylans, in Xylans and Xylanases, Visser *et al.*, eds., Elsevier, New York, N.Y., pp. 203-212; Coughlan and Hazlewood, 1993, Hemicellulose and hemicellulases, Portland, London, UK; Brigham *et al.*, 1996, Hemicellulases: Diversity and applications, in Handbook on Bioethanol: Production and Utilization, Wyman, ed., Taylor & Francis, Washington, D.C., pp. 119-141).

Suitable hemicellulases for use with the compositions and/or methods of the present disclosure include, for example, xylanases, arabinofuranosidases, acetyl xylan esterase, glucuronidases, endo-galactanase, mannanases, endo or exo arabinases, exo-galactanases, and mixtures thereof. Examples of endo-acting hemicellulases and ancillary enzymes include, without limitation, endoarabinanase, endoarabinogalactanase, endoglucanase, endomannanase, endoxylanase, and feraxan

endoxylanase. Examples of exo-acting hemicellulases and ancillary enzymes include, without limitation, α-L-arabinosidase, β-L-arabinosidase, α-1,2-L-fucosidase, α-D-galactosidase, β-D-galactosidase, β-D-glucosidase, β-D-glucuronidase, β-D-mannosidase, β-D-xylosidase, exoglucosidase, exocellobiohydrolase, exomannobiohydrolase, exomannanase, exoxylanase, xylan α-glucuronidase, and coniferin β-glucosidase. Examples of esterases include, without limitation, acetyl esterases (acetylgalactan esterase, acetylmannan esterase, and acetylxylan esterase) and aryl esterases (coumaric acid esterase and ferulic acid esterase).

In certain aspects, the hemicellulase is an exo-acting hemicellulase. Preferably, the exo-acting hemicellulase has the ability to hydrolyze hemicellulose under acidic conditions, for example, at or below pH 7.

In certain aspects, the hemicellulase is added in an effective amount. For example, the hemicellulase is added to the multienzyme blends of the present disclosure in an amount of about 0.001 wt.% or more, about 0.002 wt.% or more, about 0.0025 wt.% or more, about 0.005 wt.% or more, or about 0.01 wt.% or more relative to the weight of solids in the complete fermentation medium. In another example, the hemicellulase is added to the multienzyme blends of the present disclosure in an amount of about 0.001 wt.% to about 5.0 wt.%, for example, about 0.025 wt.% to about 4.0 wt.%, about 0.005 wt.% to about 2.0 wt.% relative to the weight of solids in the complete fermentation medium.

6.3.2.1. Xylanases

The enzyme blends of the disclosure optionally comprise one or more xylanases. The term "xylanase" as used herein refers to any xylanase classified in or under EC 3.2.1.8. Suitable xylanases include, for example, a *Caldocellum saccharolyticum* xylanase (Luthi *et al.*, 1990, Appl. Environ. Microbiol. 56(9):2677-2683), a *Thermatoga maritima* xylanase (Winterhalter & Liebel, 1995, Appl. Environ. Microbiol. 61(5):1810–1815), a *Thermatoga Sp.* Strain FJSS-B.1 xylanase (Simpson *et al.*, 1991, Biochem. J. 277, 413-417), a *Bacillus circulans* xylanase (BcX) (U.S. Patent No. 5,405,769), an *Aspergillus niger* xylanase (Kinoshita *et al.*, 1995, J. Ferment. Bioeng. 79(5):422-428); a *Streptomyces 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); *Bacillus subtilis* xylanase (Bernier *et al.*, 1983, Gene 26(1):59–65); a *Cellulomonas fimi* xylanase (Clarke *et al.*, 1996, FEMS Microbiol. Lett. 139:27-35), a *Pseudomonas fluorescens* xylanase (Gilbert *et al.*, 1988, J. Gen. Microbiol. 134:3239-3247); a *Clostridium thermocellum* xylanase (Dominguez *et al.*,1995, Nat. Struct. Biol. 2(7):569-76); a *Bacillus*

pumilus xylanase (Nuyens *et al.*, 2001, Appl. Microbiol. Biotech. 56:431-434; Yang *et al.*,1988, Nucleic Acids Res. 16(14B):7187); a *Clostridium acetobutylicum* P262 xylanase (Zappe *et al.*, 1990, Nucleic Acids Res. 18(8):2179) or a *Trichoderma harzianum* xylanase (Rose *et al.*, 1987, J. Mol. Biol. 194(4):755–756).

Xylanases can suitably be obtained from a number of sources, including, for example, fungal and bacterial organisms, such as *Aspergillus*, *Disporotrichum*, *Penicillium*, *Neurospora*, *Fusarium*, *Trichoderma*, Humicola, *Thermomyces*, and *Bacillus*. Certain commercially available preparations comprising xylanase(s) can also be used in the compositions and methods of the present disclosure; those include Multifect® xylanase, Laminex® BG and Spezyme® CP (Danisco US, Genencor), and Celluclast® and Viscozyme® (Novozymes A/S).

In certain aspects, the xylanase does not have retaining β-xylosidase activity and/or inverting β-xylosidase activity. An enzyme can be tested for retaining vs. inverting activity as described in Section 6.3.1.5 above.

6.3.2.2. β-Xylosidases

The enzyme blends of the disclosure optionally comprise one or more β-xylosidases.

As used herein, the term "β-xylosidase" refers to any β-xylosidase classified in or under EC 3.2.1.37. Suitable β-xylosidases include, for example *Talaromyces emersonii* BxI1 (Reen *et al.*, 2003, Biochem. Biophys. Res. Commun. 305(3):579-85); as well as β-xylosidases obtained from *Geobacillus stearothermophilus* (Shallom *et al.*, 2005, Biochem. 44:387-397); *Scytalidium thermophilum* (Zanoelo *et al.*, 2004, J. Ind. Microbiol. Biotechnol. 31:170-176); *Trichoderma lignorum* (Schmidt, 1988, Methods Enzymol. 160:662-671); *Aspergillus awamori* (Kurakake *et al.*, 2005, Biochim. Biophys. Acta 1726:272-279); *Aspergillus versicolor* (Andrade *et al.*, Process Biochem. 39:1931-1938); *Streptomyces sp.* (Pinphanichakarn *et al.*, 2004, World J. Microbiol. Biotechnol. 20:727-733); *Thermotoga maritima* (Xue and Shao, 2004, Biotechnol. Lett. 26:1511-1515); *Trichoderma sp.* SY (Kim *et al.*, 2004, J. Microbiol. Biotechnol. 14:643-645); *Aspergillus niger* (Oguntimein and Reilly, 1980, Biotechnol. Bioeng. 22:1143-1154); or *Penicillium wortmanni* (Matsuo *et al.*, 1987, Agric. Biol. Chem. 51:2367-2379).

In certain aspects, the β-xylosidase does not have retaining β-xylosidase activity. In other aspects, the β-xylosidase has inverting β-xylosidase activity. In yet further aspects, the β-xylosidase has no retaining β-xylosidase activity but has inverting β-xylosidase activity. An enzyme can be tested for retaining vs. inverting activity as described in Section 6.3.1.5 above.

6.3.2.3. L-α-Arabinofuranosidases

The enzyme blends of the disclosure optionally comprise one or more L-α-arabinofuranosidases.

As used herein, the term "L-α-arabinofuranosidase" refers to any enzyme classified in or under EC 3.2.1.55. Suitable L-α-arabinofuranosidase can be obtained from, for example, *Aspergillus oryzae* (Numan & Bhosle, 2006, J. Ind. Microbiol. Biotechnol. 33:247-260); *Aspergillus sojae* (Oshima *et al.*, 2005, J. Appl. Glycosci. 52:261-265); *Bacillus brevis* (Numan & Bhosle, 2006, J. Ind. Microbiol. Biotechnol. 33:247-260); *Bacillus stearothermophilus* (Kim *et al.*, 2004, J. Microbiol. Biotechnol. 14:474-482); *Bifidobacterium breve* (Shin *et al.*, 2003, Appl. Environ. Microbiol. 69:7116-7123; *Bifidobacterium longum* (Margolles *et al.*, 2003, Appl. Environ. Microbiol. 69:5096-5103); *Clostridium thermocellum* (Taylor *et al.*, 2006, Biochem. J. 395:31-37); *Fusarium oxysporum* (Panagiotou *et al.*, 2003, Can. J. Microbiol. 49:639-644); *Fusarium oxysporum f. sp. dianthi* (Numan & Bhosle, 2006, J. Ind. Microbiol. Biotechnol. 33:247-260); *Geobacillus stearothermophilus* T-6 (Shallom *et al.*, 2002, J. Biol. Chem. 277:43667-43673); *Hordeum vulgare* (Lee *et al.*, 2003, J. Biol. Chem. 278:5377-5387); *Penicillium chrysogenum* (Sakamoto *et al.*, 2003, Biophys. Acta 1621:204-210); *Penicillium sp.* (Rahman *et al.*, 2003, Can. J. Microbiol. 49:58-64); *Pseudomonas cellulosa* (Numan & Bhosle, 2006, J. Ind. Microbiol. Biotechnol. 33:247-260); *Rhizomucor pusillus* (Rahman *et al.*, 2003, Carbohydr. Res. 338:1469-1476) ; *Streptomyces chartreusis* (Numan & Bhosle, 2006, J. Ind. Microbiol. Biotechnol. 33:247-260); *Streptomyces thermoviolacus* (Numan & Bhosle, 2006, J. Ind. Microbiol. Biotechnol. 33:247-260); *Thermoanaerobacter ethanolicus* (Numan & Bhosle, 2006, J. Ind. Microbiol. Biotechnol. 33:247-260); *Thermobacillus xylanilyticus* (Numan & Bhosle, 2006, J. Ind. Microbiol. Biotechnol. 33:247-260); *Thermomonospora fusca* (Tuncer and Ball, 2003, Folia Microbiol. (Praha) 48:168-172); *Thermotoga maritima* (Miyazaki, 2005, Extremophiles 9:399-406); *Trichoderma sp.* SY (Jung *et al.*, 2005, Agric. Chem. Biotechnol. 48:7-10); *Aspergillus kawachii* (Koseki *et al.*, 2006, Biochim. Biophys. Acta 1760:1458-1464); *Fusarium oxysporum f. sp. dianthi* (Chacon-Martinez *et al.*, 2004, Physiol. Mol. Plant Pathol. 64:201-208); *Thermobacillus xylanilyticus* (Debeche *et al.*, 2002, Protein Eng. 15:21-28); *Humicola insolens* (Sorensen *et al.*, 2007, Biotechnol. Prog. 23:100-107); *Meripilus giganteus* (Sorensen *et al.*, 2007, Biotechnol. Prog. 23:100-107); or *Raphanus sativus* (Kotake *et al.*, 2006, J. Exp. Bot. 57:2353-2362).

In certain aspects, the L-α-arabinofuranosidase does not have retaining β-xylosidase activity. In other aspects, the L-α-arabinofuranosidase has inverting β-

xylosidase activity. In yet further aspects, the L-α-arabinofuranosidase has no retaining β-xylosidase but has inverting β-xylosidase activity. An enzyme can be tested for retaining vs. inverting activity as described in Section 6.3.1.5 above.

6.3.3. Accessory proteins

A number of polypeptides having cellulolytic enhancing activity can also be used in conjunction with the above-noted enzymes and/or cellulolytic proteins to further degrade the cellulose component of the biomass substrate, (*see, e.g.*, Brigham *et al.*, 1995, in Handbook on Bioethanol (Charles E. Wyman, ed.), pp. 119-141, Taylor & Francis, Washington D.C.; Lee, 1997, J. Biotechnol. 56: 1-24).

The optimum amounts of such a polypeptide having cellulolytic enhancing activity and of cellulolytic proteins depend on a number of factors including, without limitation, the specific mixture of component cellulolytic proteins, the cellulosic substrate, the concentration of cellulosic substrate, the pretreatment(s) of the cellulosic substrate, the temperature, time, and pH, and the nature of the fermenting organism.

The enzyme blends/compositions of the disclosure can, for example, suitably further comprise one 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, glycoside hydrolase Family 61 polypeptides, xyloglucanases, CIP1, CIP2, swollenin, expansins, and cellulose disrupting proteins. Examples of accessory proteins can also include CIP1-like proteins, CIP2-like proteins, cellobiose dehydrogenases and manganese peroxidases. In particular embodiments, the cellulose disrupting proteins are cellulose binding modules.

6.4 Enzymes with Inverting β-xylosidase Activity

According to the present disclosure, an enzyme with inverting β-xylosidase activity is used to reduce AXP (e.g., EXP) formation in SSF reactions. Thus, the present disclosure pertains, in one aspect, to a composition comprising at least one inverting β-xylosidase polypeptide. In another aspect, the present disclosure pertains to a method of producing a desired fermentation product in an SSF reaction comprising culturing a complete fermentation medium, said complete fermentation medium comprises at least one inverting β-xylosidase polypeptide.

Suitable inverting β-xylosidase polypeptides can be selected from those that are members of the glycoside hydrolase family 43 ("GH43"). GH43 family enzymes have a

number of known activities. For example, a GH43 family enzyme can be one that is classified under EC 3.2.1.55 and can have L-α-arabinofuranosidase activity. In another example, a GH43 family enzyme can be one that is classified under EC 3.2.1.99, and can have endo-arabinanase activity. In yet another example, a GH43 family enzyme can be classified under EC 3.2.1.145, and can have galactan 1,3-β-galactosidase activity. In other examples, GH43 family of enzymes can be classified under EC 3.2.1.37 and can have β-xylosidase activity. Whilst GH43 family of β-xylosidases, such as those described above, often can only perform inverting hydrolysis, various β-xylosidases from the GH3, -39, -52, and -54 families, in contrast, have been reported to have retaining activities and to be able to perform both hydrolysis and transglycosylation reactions. (Smaali *et al.*, 2006, Appl. Microbiol. Biotechnol. 73:582-590).

GH43 family enzymes typically display a five-bladed-β-propeller three-dimensional conformation. The "propeller" part of the structure is based upon a five-fold repeat of "blade"-like conformation that comprisesfour-stranded β-sheets. The catalytic general base, an aspartate, the catalytic general acid, a glutamate, and an aspartate residue that modulates the pKa of the general base have been identified through the crystal structure of *Cellvibrio japonicus* CjArb43A, and have been confirmed by site-directed mutagenesis (*see* Nurizzo *et al.*, 2002, Nat. Struct. Biol. 9(9) 665-8). The catalytic residues are arranged in three conserved blocks, which spread widely throughout the amino acid sequence (Pons *et al.*, 2004, Proteins: Structure, Function and Bioinformatics 54:424-432). For GH43 family of β-xylosidase enzymes, the predicted catalytic residues are shown as the bold and underlined type face fonts in the sequences of Figure 32. The crystal structure of *Geobacillus stearothermophilus* xylosidase (Brux *et al.* 2006, J. Mol. Bio. 359:97-109) suggests several additional residues that might be important for substrate binding in that enzyme.

As described in Section 6.3.1.5 above, inverting β-xylosidase activity can be determined by suitable assays.

Accordingly, in certain aspects, the enzyme with inverting β-xylosidase activity herein is a GH43 family member. For example, the enzyme is an Fv43D, a Pf43A, an Fv43E, an Fv43B, an Af43A, an Fo43A, a Gz43A, or a XynB3 polypeptide. Such polypeptides are described below.

6.4.1. Fv43D Polypeptides

In certain embodiments, the enzyme with inverting β-xylosidase activity is an Fv43D polypeptide. The amino acid sequence of Fv43D (SEQ ID NO:2) is shown in Figure 19B and on the first line of Figure 32. SEQ ID NO:2 is the sequence of the

immature Fv43D. Fv43D has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:2 (underlined in Figure 19B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 21 to 350 of SEQ ID NO:2. The predicted conserved domain residues are in boldface type in Figure 19B. Fv43D was shown to have β-xylosidase activity in an assay using *p*-nitrophenyl-β-xylopyranoside, xylobiose, or mixed, linear xylo-oligomers as substrates. The predicted catalytic residues are: either D37 or D71; D155; and E251.

As used herein, "an Fv43D polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.,* at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.,* at least 75, 100, 125, 150, 175, 200, 250, 300, or 320 contiguous amino acid residues among residues 21 to 350 of SEQ ID NO:2. An Fv43D polypeptide preferably is unaltered as compared to native Fv43D in residues D37 or D71; D155, and E251. An Fv43D polypeptide is preferably unaltered in at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, or all nine of Fv43D, Fo43A, Gz43A, Pf43A, Fv43A, Fv43B, Af43A, Pf43B, and Fv43E, as shown in the alignment of Figure 32. An Fv43D polypeptide suitably comprises the entire predicted conserved domain of native Fv43D as shown in Figure 19B. An exemplary Fv43D polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Fv43D sequence as shown in Figure 19B. The Fv43D polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Fv43D polypeptide of the invention has inverting β-xylosidase activity.

6.4.2. Pf43A Polypeptides

In certain embodiments, the enzyme with inverting β-xylosidase activity is a Pf43A polypeptide. The amino acid sequence of Pf43A (SEQ ID NO:8) is shown in Figure 22B and on the fourth line of Figure 32. SEQ ID NO:8 is the sequence of the immature Pf43A. Pf43A has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:8 (underlined in Figure 22B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 21 to 445 of SEQ ID NO:8. The predicted catalytic domain residues are in boldface type, the predicted carbohydrate binding domain residues are in uppercase type, and the predicted linker residues separating the catalytic domain and carbohydrate binding

domain are in italics in Figure 22B. Pf43A was shown to have β-xylosidase activity in an assay using p-nitrophenyl-β-xylopyranoside, xylobiose or mixed, linear xylo-oligomers as substrates. The predicted catalytic residues are: either D32 or D60; D145; and E196.

As used herein, "a Pf43A polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.,* at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.,* at least 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 contiguous amino acid residues among residues 21 to 445 of SEQ ID NO:8. A Pf43A polypeptide preferably is unaltered as compared to native Pf43A in residues D32 or D60; D145, and E196. A Pf43A polypeptide is preferably unaltered in at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, or all nine of Fv43D, Fo43A, Gz43A, Pf43A, Fv43A, Fv43B, Af43A, Pf43B, and Fv43E, as shown in the alignment of Figure 32. A Pf43A polypeptide suitably comprises the entire predicted conserved domain of native Pf43A as shown in Figure 22B. An exemplary Pf43A polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Pf43A sequence as shown in Figure 22B. The Pf43A polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Pf43A polypeptide of the invention has inverting β-xylosidase activity.

6.4.3. Fv43E Polypeptides

In certain embodiments, the enzyme with inverting β-xylosidase activity is an Fv43E polypeptide. The amino acid sequence of Fv43E (SEQ ID NO:10) is shown in Figure 23B and on the ninth line of Figure 32. SEQ ID NO:10 is the sequence of the immature Fv43E. Fv43E has a predicted signal sequence corresponding to residues 1 to 18 of SEQ ID NO:10 (underlined in Figure 23B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 19 to 530 of SEQ ID NO:10. The predicted catalytic domain residues are in boldface type in Figure 23B. Fv43E was shown to have β-xylosidase activity in an assay using p-nitrophenyl-β-xylopyranoside, xylobiose and mixed, linear xylo-oligomers as substrates. The predicted catalytic residues are: either D40 or D71; D155; and E242.

As used herein, "an Fv43E polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.,* at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence

identity to at least 50, *e.g.,* at least 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, or 500 contiguous amino acid residues among residues 19 to 530 of SEQ ID NO:10. An Fv43E polypeptide preferably is unaltered as compared to native Fv43E in residues D40 or D71; D155; and E242. An Fv43E polypeptide is preferably unaltered in at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, or all nine of Fv43D, Fo43A, Gz43A, Pf43A, Fv43A, Fv43B, Af43A, Pf43B, and Fv43E, as shown in the alignment of Figure 32. An Fv43E polypeptide suitably comprises the entire predicted conserved domain of native Fv43E as shown in Figure 23B. An exemplary Fv43E polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Fv43E sequence as shown in Figure 23B. The Fv43E polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Fv43E polypeptide of the invention has inverting β-xylosidase activity.

6.4.4. Fv43B Polypeptides

In certain embodiments, the enzyme with inverting β-xylosidase activity is an Fv43B polypeptide. The amino acid sequence of Fv43B (SEQ ID NO:12) is shown in Figure 24B and on the sixth line of Figure 32. SEQ ID NO:12 is the sequence of the immature Fv43B. Fv43B has a predicted signal sequence corresponding to residues 1 to 16 of SEQ ID NO:12 (underlined in Figure 24B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 17 to 574 of SEQ ID NO:12. The predicted catalytic domain residues are in boldface type in Figure 24B. Fv43B was shown to have both β-xylosidase and L-α-arabinofuranosidase activity in assays using *p*-nitrophenyl-β-xylopyranoside and/or *p*-nitrophenyl-α-L-arabinofuranoside as substrates. It was shown to release arabinose from branched arabino-xylooligomers and to increase xylose release from oligomer mixtures in the presence of other xylosidase enzymes. The predicted catalytic residues are: either D38 or D68; D151; and E236.

As used herein, "an Fv43B polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.,* at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.,* at least 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, or 550 contiguous amino acid residues among residues 17 to 472 of SEQ ID NO:12. An Fv43B polypeptide preferably is unaltered as compared to native Fv43B in

residues D38 or D68; D151; and E236. An Fv43B polypeptide is preferably unaltered in at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, or all nine of Fv43D, Fo43A, Gz43A, Pf43A, Fv43A, Fv43B, Af43A, Pf43B, and Fv43E, as shown in the alignment of Figure 32. An Fv43B polypeptide suitably comprises the entire predicted conserved domain of native Fv43B as shown in Figure 24B. An exemplary Fv43B polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Fv43B sequence as shown in Figure 24B. The Fv43B polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Fv43B polypeptide of the invention has inverting β-xylosidase activity.

6.4.5. Af43A Polypeptides

In certain embodiments, the enzyme with inverting β-xylosidase activity is an Af43A polypeptide. The amino acid sequence of Af43A (SEQ ID NO:14) is shown in Figure 25B and on the seventh line of Figure 32. SEQ ID NO:14 is the sequence of the immature Af43A. The predicted conserved domain residues are in boldface type in Figure 25B. Af43A was shown to have L-α-arabinofuranosidase activity in an assay using *p*-nitrophenyl-α-L-arabinofuranoside and by the release of arabinose from converting the set of oligomers produced *via* the action of an endoxylanase. The predicted catalytic residues are: either D26 or D58; D139; and E227.

As used herein, "an Af43A polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.,* at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.,* at least 75, 100, 125, 150, 175, 200, 250, or 300 contiguous amino acid residues of SEQ ID NO:14. An Af43A polypeptide preferably is unaltered as compared to native Af43A in residues D26 or D58; D139; and E227. An Af43A polypeptide is preferably unaltered in at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, or all nine of Fv43D, Fo43A, Gz43A, Pf43A, Fv43A, Fv43B, Af43A, Pf43B, and Fv43E, as shown in the alignment of Figure 32. An Af43A polypeptide suitably comprises the entire predicted conserved domain of native Af43A as shown in Figure 25B. An exemplary Fv43B polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%

identity to the mature Af43A sequence as shown in Figure 25B. The Af43A polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Af43A polypeptide of the invention has inverting β-xylosidase activity.

6.4.6. Fo43A Polypeptides

In certain embodiments, the enzyme with inverting β-xylosidase activity is an Fo43A polypeptide. The amino acid sequence of Fo43A (SEQ ID NO:24) is shown in Figure 31B and on the second line of Figure 32. SEQ ID NO:24 is the sequence of the immature Fo43A. Fo43A has a predicted signal sequence corresponding to residues 1 to 17 of SEQ ID NO:24 (underlined in Figure 31B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 21 to 348 of SEQ ID NO:24. The predicted conserved domain residues are in boldface type in Figure 31B. Fo43A was shown to have β-xylosidase activity in an assay using p-nitrophenyl-β-xylopyranoside, xylobiose or mixed, linear xylo-oligomers as substrates. The predicted catalytic residues are: either D37 or D72; D159; and E251.

As used herein, "an Fo43A polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.*, at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.*, at least 75, 100, 125, 150, 175, 200, 250, 300, or 320 contiguous amino acid residues among residues 21 to 348 of SEQ ID NO:24. An Fo43A polypeptide preferably is unaltered as compared to native Fo43A in residues D37 or D72; D159; and E251. An Fo43A polypeptide is preferably unaltered in at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, or all nine of Fv43D, Fo43A, Gz43A, Pf43A, Fv43A, Fv43B, Af43A, Pf43B, and Fv43E, as shown in the alignment of Figure 32. An Fo43A polypeptide suitably comprises the entire predicted conserved domain of native Fo43A as shown in Figure 31B. An exemplary Fo43A polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Fo43A sequence as shown in Figure 31B. The Fo43A polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Fo43A polypeptide of the invention has inverting β-xylosidase activity.

6.4.7. Gz43A Polypeptides

In certain embodiments, the enzyme with inverting β-xylosidase activity is a Gz43A polypeptide. The amino acid sequence of Gz43A (SEQ ID NO:22) is shown in

Figure 30B and on the third line of Figure 32. SEQ ID NO:22 is the sequence of the immature Gz43A. Gz43A has a predicted signal sequence corresponding to residuess 1 to 18 of SEQ ID NO:22 (underlined in Figure 30B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 19 to 340 of SEQ ID NO:22. The predicted conserved domain residues are in boldface type in Figure 30B. Gz43A was shown to have β-xylosidase activity in an assay using p-nitrophenyl-β-xylopyranoside, xylobiose or mixed, linear xylo-oligomers as substrates. The predicted catalytic residues are: either D33 or D68; D154; and E243.

As used herein, "a Gz43A polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.,* at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.,* at least 75, 100, 125, 150, 175, 200, 250, or 300 contiguous amino acid residues among residues 19 to 340 of SEQ ID NO:22. A Gz43A polypeptide preferably is unaltered as compared to native Gz43A in residues either D33 or D68; D154; and E243. A Gz43A polypeptide is preferably unaltered in at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, or all nine of Fv43D, Fo43A, Gz43A, Pf43A, Fv43A, Fv43B, Af43A, Pf43B, and Fv43E, as shown in the alignment of Figure 32. A Gz43A polypeptide suitably comprises the entire predicted conserved domain of native Gz43A as shown in Figure 30B. An exemplary Gz43A polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Fo43A sequence as shown in Figure 30B. The Gz43A polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Gz43A polypeptide of the invention has inverting β-xylosidase activity.

6.4.8. *G. stearothermophilus* XynB3 Polypeptides

In other aspects, the enzyme with inverting β-xylosidase activity is a *G. stearothermophilus* XynB3 polypeptide. The sequence of *G. stearothermophilus* XynB3 is presented as SEQ ID NO:25. *G. stearothermophilus* XynB3 is a 535-amino-acid GH43 family enzymefrom *Geobacillus stearothermophilus* T-6. The enzyme cleaves single xylose units from the non-reducing end of xylooligomers; the three catalytic residues D15, D128, and E187 were found to be essential for its activity (Shallom *et al.*, 2005, Biochemistry, 44:387–397).

As used herein, "a *G. stearothermophilus* XynB3 polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.*, at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.*, at least 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, or 500 contiguous amino acid residues of SEQ ID NO:25. A *G. stearothermophilus* XynB3 polypeptide preferably is unaltered as compared to native XynB3 in residues D15, D128, and E187. The *G. stearothermophilus* XynB3 polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the *G. stearothermophilus* XynB3 polypeptide of the invention has inverting β-xylosidase activity.

6.4.9. *Vibrio sp.* XloA Polypeptides

In certain embodiments, the enyme with inverting β-xylosidase activity is a *Vibrio sp.* XloA polypeptide. *Vibrio sp.* XloA is a β-1,3-xylosidase from *Vibrio sp.* strain XY-214.

As used herein, "a *Vibrio sp.* XloA polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.*, at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.*, at least 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 contiguous amino acid residues of β-1,3-Xylosidase from *Vibrio sp.* Strain XY-214 (Umemoto *et al.*, 2008, Appl. Environ. Microbiol. 74(1): 305–308; Genbank Accession No. AB300564). The *Vibrio sp.* XloA polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the *Vibrio sp.* XloA polypeptide of the invention has inverting β-xylosidase activity.

6.5 Enzymes with Retaining β-xylosidase Activity

According to the present disclosure, an enzyme with retaining β-xylosidase activity is used to improve or increase AXP (e.g., EXP) production in SSF reactions. Thus, the present disclosure pertains, in one aspect, to a composition comprising at least one retaining β-xylosidase polypeptide. In another aspect, the present disclosure pertains to a method of producing a desired AXP compound or an improved or increased amount of an AXP product in an SSF reaction comprising culturing a complete fermentation medium, said complete fermentation medium comprises at least one retaining β-xylosidase polypeptide.

Suitable inverting β-xylosidase polypeptides can be selected from those that are members of the glycoside hydrolase family 3 ("GH3"), GH30, GH31, GH39, GH52, GH54, or GH116 family enzymes. .

As described in Section 6.3.1.5 above, retaining β-xylosidase activity can be determined by suitable assays.

Accordingly, in certain aspects, the enzyme with retaining β-xylosidase activity herein is a GH3, GH30, GH31, GH39, GH52, GH54, or GH116 family member. For example, the enzyme is an *Aspergillus japonicus* XlnD, a *Fusarium verticillioides* Fv30A, a *Fusarium verticillioides* Fv30B, a *Fusarium verticillioides* Fv39A, a *Fusarium verticillioides* Fv39B, a *Thermoanaerobacter saccharolyticum* XynB, a *Geobacillus stearothermophilus* XylA, or a *Trichoderma koningii (Hypocrea koningii)* Xyl1 polypeptide. Such polypeptides are described below.

6.5.1. *Aspergillus japonicus* XlnD Polypeptides

In certain embodiments, the enzyme with retaining β-xylosidase activity is a XlnD polypeptide. The amino acid sequence of XlnD (SEQ ID NO:40) is shown in Figure 35B. SEQ ID NO:40 is the sequence of the immature XlnD. XlnD has a predicted signal sequence corresponding to residuess 1 to 17 of SEQ ID NO:40 (underlined in Figure 35B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 18 to 804 of SEQ ID NO:40.

As used herein, "a XlnD polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.*, at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.*, at least 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, or 750 contiguous amino acid residues among residues 18-804 of SEQ ID NO:40. An exemplary XlnD polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature XlnD sequence as shown in Figure 35B. The XlnB polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the XlnB polypeptide of the invention has retaining β-xylosidase activity.

6.5.2. *Fusarium verticillioides* Fv30A Polypeptides

In certain embodiments, the enzyme with retaining β-xylosidase activity is an Fv30A polypeptide. The amino acid sequence of XlnD (SEQ ID NO:42) is shown in Figure 36B. SEQ ID NO:42 is the sequence of the immature Fv30A. XlnD has a predicted signal sequence corresponding to residuess 1 to 19 of SEQ ID NO:42 (underlined in Figure 36B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 20 to 537 of SEQ ID NO:42.

As used herein, "an Fv30A polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.,* at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.,* at least 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, or 500 contiguous amino acid residues among residues 20 to 537 of SEQ ID NO:42. An exemplary Fv30A polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Fv30A sequence as shown in Figure 36B. The Fv30A polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Fv30A polypeptide of the invention has retaining β-xylosidase activity.

6.5.3. *Fusarium verticillioides* Fv30B Polypeptides

In certain embodiments, the enzyme with retaining β-xylosidase activity is an Fv30B polypeptide. The amino acid sequence of Fv30B (SEQ ID NO:44) is shown in Figure 37B. SEQ ID NO:44 is the sequence of the immature Fv30B. Fv30B has a predicted signal sequence corresponding to residuess 1 to 24 of SEQ ID NO:44 (underlined in Figure 37B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 25 to 485 of SEQ ID NO:44.

As used herein, "an Fv30B polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.,* at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.,* at least 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, or 450 contiguous amino acid residues among residues 25-485 of SEQ ID NO:44. An exemplary Fv30B polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Fv30B sequence as shown in Figure 37B. The Fv30B polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Fv30B polypeptide of the invention has retaining β-xylosidase activity.

6.5.4. *Fusarium verticillioides* Fv39A Polypeptides

In certain embodiments, the enzyme with retaining β-xylosidase activity is an Fv39A polypeptide. The amino acid sequence of Fv39A (SEQ ID NO:46) is shown in Figure 38B. SEQ ID NO:46 is the sequence of the immature Fv39A. Fv39A has a predicted signal sequence corresponding to residuess 1 to 19 of SEQ ID NO:46 (underlined in Figure 38B); cleavage of the signal sequence is predicted to yield a

mature protein having a sequence corresponding to residues 20 to 439 of SEQ ID NO:46.

As used herein, "an Fv39A polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.,* at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.,* at least 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 contiguous amino acid residues among residues 20-439 of SEQ ID NO:46. An exemplary Fv39A polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Fv39A sequence as shown in Figure 38B. The Fv39A polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Fv39A polypeptide of the invention has retaining β-xylosidase activity.

6.5.5. *Fusarium verticillioides* Fv39B Polypeptides

In certain embodiments, the enzyme with retaining β-xylosidase activity is an Fv39B polypeptide. The amino acid sequence of Fv39B (SEQ ID NO:48) is shown in Figure 39B. SEQ ID NO:48 is the sequence of the immature Fv39B. Fv39B has a predicted signal sequence corresponding to residuess 1 to 18 of SEQ ID NO:48 (underlined in Figure 39B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 19 to 456 of SEQ ID NO:48.

As used herein, "an Fv39B polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.,* at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.,* at least 75, 100, 125, 150, 175, 200, 250, 300, or 350 contiguous amino acid residues among residues 19-456 of SEQ ID NO:48. An exemplary Fv39B polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Fv39B sequence as shown in Figure 39B. The Fv39B polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Fv39B polypeptide of the invention has retaining β-xylosidase activity.

6.5.6. *Thermoanaerobacter saccharolyticum* XynB Polypeptides

In certain embodiments, the enzyme with retaining β-xylosidase activity is a XynB polypeptide. The amino acid sequence of XynB (SEQ ID NO:50) is shown in Figure 40B. XynB does not have a predicted signal sequence from the SignalP algorithm (available at: http://www.cbs.dtu.dk).

As used herein, "a XynB polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.*, at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.*, at least 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, or 450 contiguous amino acid residues of SEQ ID NO:50 . An exemplary XynB polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to XynB sequence as shown in Figure 40B. The XynB polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the XynB polypeptide of the invention has retaining β-xylosidase activity.

6.5.7. *Geobacillus stearothermophilus* XylA Polypeptides

In certain embodiments, the enzyme with retaining β-xylosidase activity is a XylA polypeptide. The amino acid sequence of XylA (SEQ ID NO:52) is shown in Figure 41B. XylA does not have a predicted signal sequence from the SignalP algorithm (available at: http://www.cbs.dtu.dk), but has a signal sequence predicted from the Uniprot algorithm (available at: http://www.uniprot.org/uniprot) that corresponds to residues 1 to 18 of SEQ ID NO:52 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 19-705 of SEQ ID NO:52.

As used herein, "a XylA polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.*, at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.*, at least 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, or 650 contiguous amino acid residues of SEQ ID NO:52, or to residues 19-705 of SEQ ID NO:52. An exemplary XylA polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature XylA sequence as shown in Figure 41B. The XylA polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the XylA polypeptide of the invention has retaining β-xylosidase activity.

6.5.8. *Trichoderma koningii (Hypocrea koningii)* Xyl1 Polypeptides

In certain embodiments, the enzyme with retaining β-xylosidase activity is a Xyl1 polypeptide. The amino acid sequence of Xyl1(SEQ ID NO:54) is shown in Figure 42B. SEQ ID NO:54 is the sequence of the immature Xyl1. Xyl1 has a predicted signal sequence corresponding to residuess 1 to 21 of SEQ ID NO:54 (underlined in Figure 42B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 22 to 500 of SEQ ID NO:54.

As used herein, "a Xyl1 polypeptide" refers to a polypeptide and/or to a variant thereof comprising a sequence having at least 85%, *e.g.*, at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, *e.g.*, at least 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, or 450 contiguous amino acid residues among residues 22-500 of SEQ ID NO:54. An exemplary Xyl1 polypeptide of the invention comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to the mature Xyl1 sequence as shown in Figure 42B. The Xyl1 polypeptide of the invention suitably has β-xylosidase activity. In certain embodiments, the Xyl1 polypeptide of the invention has retaining β-xylosidase activity.

6.6 Recombinant Methods for Production of Enzymes for Use In SSF

6.6.1. Nucleic Acids and Expression Vectors

Natural or synthetic polynucleotide fragments encoding an enzyme for use in SSF ("SSF enzyme"), including an inverting β-xylosidase polypeptide or other saccharifying enzyme(s) (*e.g.*, a cellulase or a hemicellulase), can be incorporated into heterologous nucleic acid constructs or vectors. Those vectors can then be introduced into, or replicated in a suitable host cell, including, for example, a filamentous fungal, yeast, or bacterial cell. The vectors and methods disclosed herein can be used to express one or more SSF enzyme(s). Any vector can be used as long as it is replicable and viable in the cells into which it is introduced. Many suitable vectors and promoters are known to those of skill in the art, among which a large number are commercially available. Cloning and expression vectors have been extensively described in the literature, for example, in Sambrook *et al.*, 2001, Molecular Cloning: A Laboratory Manual (CSHL Press) and in Ausubel *et al.*, 2002, Short Protocols in Molecular Biology (Current Protocols), the content of each concerning expression vectors is expressly incorporated by reference herein. Other exemplary expression vectors that are suitable for fungal host cellsare described in van den Hondel *et al.*,1991, Bennett and Lasure (eds.) More Gene Manipulations in Fungi. Academic Press, pp. 396-428.

It is known in the art that various DNA sequences of interest can be inserted into plasmids or vectors (collectively referred to herein as "vectors") using a number of standard procedures. Typically, for example, the DNA sequence of interest is inserted into an appropriate restriction endonuclease site using standard procedures and under standard conditions. Such procedures and related sub-cloning procedures are \ within the scope of knowledge of those ordinarily skilled in the art.

Recombinant filamentous fungi comprising the coding sequence for an SSF enzyme can be produced by introducing a heterologous nucleic acid construct comprising the SSF enzyme coding DNA sequence into the genetic material of the filamentous fungi host cells.

Once the desired form of nucleic acid sequence encoding the SSF enzyme is obtained, it can optionally be modified in a variety of ways. Where the sequence involves non-coding flanking regions, the flanking regions can be subjected to resection, mutagenesis, *etc.* Thus, transitions, transversions, deletions, and/or insertions can be performed on the naturally occurring sequence.

A selected SSF enzyme coding sequence can be inserted into a suitable vector according to well-known recombinant techniques, which can then be used to transform a filamentous fungal host cell capable of expressing the SSF enzyme. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence can be used to clone and express the SSF enzyme.

The present disclosure also includes recombinant nucleic acid constructs comprising one or more of the SSF enzyme-encoding nucleic acid sequences as described above. The constructs each suitably comprises a vector, such as a plasmid or a viral vector, into which a sequence of the disclosure has been inserted, in a forward or reverse orientation.

A heterologous nucleic acid construct can suitably include the coding sequence of an SSF enzyme: (i) in isolation; (ii) in combination with additional coding sequences, such as, for example, fusion protein or signal peptide coding sequences, where the desired SSF enzyme coding sequence is the dominant coding sequence; (iii) in combination with one or more non-coding sequences, such as, for example, introns and control elements, such as, for example, promoter and terminator elements, or 5' and/or 3' untranslated regions, effective for expression of the coding sequence in a suitable host; and/or (iv) in a vector or host environment in which the SSF enzyme coding sequence is heterologous relative to the host cell.

In certain aspects, a heterologous nucleic acid construct is employed to transfer an SSF enzyme-encoding nucleic acid sequence into a cell *in vitro*, for example, a cell of an established filamentous fungal or yeast lines. For long-term production of an SSF enzyme, stable expression is preferred. It follows that any method effective to generate stable transformants can suitably be used to practice the invention disclosed herein.

Appropriate vectors are typically equipped with a selectable marker-encoding nucleic acid sequence, insertion sites, and suitable control elements, such as, for example, promoter and termination sequences. The vectors may comprise regulatory sequences, including, for example, non-coding sequences, such as introns and control elements that are operably linked to the coding sequences, and that are effective for expression of the coding sequence in the host cells. Suitable control elements include, for example, promoter and terminator elements, or 5' and/or 3' untranslated regions. A number of vectors and promoters are known to those of skill in the art, and among which, many are commercially available. Suitable vectors and promoters are also described in the literature, for example, in Sambrook, *et al.*, 2001, Molecular Cloning: A Laboratory Manual (CSHL Press).

Exemplary promoters include, for example, constitutive promoters and inducible promoters, such as, without limitation, a CMV promoter, an SV40 early promoter, an RSV promoter, an EF-1α promoter, a promoter containing the tet responsive element (TRE) in the tet-on or tet-off system (*see, e.g.,* ClonTech's description of its Tet-On® and Tet-Off® Advanced Inducible Gene Expression System), or the β actin promoter and the metallothionine promoter that can upregulated by addition of certain metal salts. A promoter sequence is a DNA sequence which is recognized by the particular filamentous fungal host cell for expression purposes. It is operably linked to DNA sequence encoding the SSF enzyme of interest. Such a linkage positions the promoter with respect to the initiation codon of the DNA sequence encoding the SSF enzyme of interest in the expression vector. The promoter sequence contains transcription and/or translation control sequences, which mediate the expression of the SSF enzyme of interest. Examples include promoters from the *Aspergillus niger*, *A. awamori*, or *A. oryzae* glucoamylase, α-amylase, or α-glucosidase encoding genes; the *A. nidulans gpdA* or *trpC* genes; the *Neurospora crassa cbh1* or *trp1* genes; the *A. niger* or *Rhizomucor miehei* aspartic proteinase encoding genes; the *H. jecorina cbh1*, *cbh2*, *egl1*, *egl2*, or other cellulase encoding genes.

The choice of selectable marker will depend on the host cell, and appropriate selectable markers suitable for use in different host cells are known in the art. Exemplary selectable marker genes include *argB* from *A. nidulans* or *H. jecorina*, *amdS* from *A. nidulans*, *pyr4* from *Neurospora crassa* or *H. jecorina*, *pyrG* from *Aspergillus niger* or *A. nidulans*. Other suitable selectable markers include, for example, *trpc, trp1, oliC31, niaD* or *leu2*, which are included in heterologous nucleic acid constructs used to transform a mutant strain such as a trp^-, pyr^-, or leu^- mutant strain, or the like.

Such selectable markers can confer, to the transformants, the ability to utilize a metabolite that is otherwise not metabolized by the host cell. For example, the *amdS* gene from *H. jecorina,* which encodes the enzyme acetamidase, allows the transformant cells to grow on acetamide as a nitrogen source. In a further example, selectable marker (*e.g.*, *pyrG*) can restore the ability of an auxotrophic mutant strain to grow on a selective minimal medium. In yet another example, selectable marker (*e.g.*, olic31) can confer, to the transformants, the ability to grow in the presence of an inhibitory drug or an antibiotic.

The selectable marker coding sequence is suitably cloned into a plasmid using methods and techniques known in the art. Exemplary plasmids include, without limitation, pUC18, pBR322, pRAX, and pUC100. For example, the pRAX plasmid contains AMAL sequences from *A. nidulans*, making it possible to replicate in *A. niger.*

The practice of the present disclosure will employ, unless otherwise specifically indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and immunology, which are within the ordinary skill in the art. Such techniques are described extensively in the literature. *See, e.g.,* Sambrook *et al.*, 2001, Moecular Cloning: A Laboratory Manual (CSHL Press); Ausubel *et al.*, 2002, Short Protocols in Molecular Biology (Current Protocols); Freshney, 2005, Culture of Animal Cells: A Manual of Basic Technique (Wiley-Liss); and Dunn *et al.*, 2003, Short Protocols in Protein Science (Wiley). All patents, patent applications, articles and publications mentioned herein, are hereby incorporated by reference.

6.6.2. Host Organisms and Protein Expression

Provided by this disclosure are host cells that are engineered to express an SSF protein of interest for use in the methods described herein. Suitable host cells include any microorganism (*e.g.*, a bacterium, a protist, an alga, a fungus (*e.g.*, a yeast, or a filamentous fungus), or any other microbe). Suitable host cell is preferably a bacterium, a yeast, or a filamentous fungus cell.

Suitable bacterial genera include, but are not limited to, *Escherichia*, *Bacillus*, *Lactobacillus*, *Pseudomonas* and *Streptomyces*. Suitable bacterial species include, but are not limited to, *Escherichia coli*, *Bacillus subtilis*, *Bacillus licheniformis*, *Lactobacillus brevis*, *Pseudomonas aeruginosa* and *Streptomyces lividans*.

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

Suitable filamentous fungi include all filamentous forms of the subdivision Eumycotina. Suitable filamentous fungal genera include, but are not limited to, *Acremonium, Aspergillus, Aureobasidium, Bjerkandera, Ceriporiopsis, Chrysoporium, Coprinus, Coriolus, Corynascus, Chaertomium, Cryptococcus, Filobasidium, Fusarium, Gibberella, Humicola, Magnaporthe, Mucor, Myceliophthora, Mucor, Neocallimastix, Neurospora, Paecilomyces, Penicillium, Phanerochaete, Phlebia, Piromyces, Pleurotus,Scytaldium, Schizophyllum, Sporotrichum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, Trametes, and Trichoderma.*

Suitable filamentous fungal species include, but are not limited to, *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 radiate, Pleurotus eryngii, Talaromyces flavus, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei*, and *Trichoderma viride.*

Once a recombinant SSF enzyme expression construct has been generated, for example, in accordance with the methods described herein, the construct can be transformed into a suitable host cell using routine methodology.

6.6.3. Methods of Enzyme Isolation and/or Purification

In certain aspects, a recombinant SSF enzyme is engineered with a signal sequence such that the recombinant SSF enzyme is secreted into the culture medium of the host cell. In certain aspects, the SSF enzyme of interest is recovered in the form of fermentation broth. The term "fermentation broth," as used herein, refers to an enzyme preparation produced by fermentation that then undergoes no or minimal recovery and/or purification thereafter. For example, microbial cultures are grown to saturation,

incubated under carbon-limiting conditions to allow protein synthesis (*e.g.*, expression of enzymes), and once the enzyme is secreted into the cell culture medium, the fermentation broth is one from which an SSF enzyme of interest can be recovered. The fermentation broth can, for example, contain the unfractionated or fractionated contents of the fermentation materials derived at the end of the fermentation. Typically, the fermentation broth is unfractionated and comprises the spent culture medium and cell debris present after the microbial cells (*e.g.*, filamentous fungal cells) are removed, *e.g.*, by centrifugation. In certain embodiments, the fermentation broth contains the spent cell culture medium, extracellular enzymes, and either live or killed microbial cells. In some embodiments, the fermentation broth is fractionated to remove the microbial cells, and as such comprises the spent cell culture medium and extracellular enzymes.

In some aspects, partial or complete purification of an SSF enzyme may be desirable. In certain embodiments, an SSF enzyme is purified to at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 95%, at least 98%, or at least 99% homogeneity.

However, in certain other aspects, an SSF enzyme of interest can be produced in a cellular form (*i.e.,* partially or entirely not secreted), which then may require recovery from a cell lysate. In such cases, the SSF enzyme is purified from the cells in which it was produced using techniques routinely employed in the art. Examples of such techniques include, without limitation, affinity chromatography (*see, e.g.,* van Tilbeurgh *et al.*, 1984, FEBS Lett. 169:215-218), ion-exchange chromatographic methods (*see, e.g.,* Goyal *et al.*, 1991, Bioresource Technol. 36:37-50; Fliess *et al.*, 1983, Eur. J. Appl. Microbiol. Biotechnol. 17:314-318; Bhikhabhai *et al.*, 1984, J. Appl. Biochem. 6:336-345; Ellouz *et al.*, 1987, J. Chromatography 396:307-317), ion-exchange chromatographic methods employing materials that have high resolution power (*see, e.g.,* Medve *et al.*, 1998, J. Chromatography A 808:153-165), hydrophobic interaction chromatography (*see, e.g.,* Tomaz and Queiroz, 1999, J. Chromatography A 865:123-128), and two-phase partitioning (*see, e.g.,* Brumbauer, *et al.*, 1999, Bioseparation 7:287-295).

Suitably, the SSF enzyme is fractionated to segregate proteins having pre-identified properties, such as binding affinity to particular binding agents or media, *e.g.*, antibodies or receptors; a certain molecular weight range; or a certain isolectric point range.

Once expression of a given SSF enzyme is achieved, the SSF enzyme thereby produced can be purified from the cells or from the cell culture. Exemplary procedures suitable for such purification include, without limitation, antibody-affinity column

chromatography; ion exchange chromatography; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; and gel filtration using, *e.g.*, Sephadex G-75. A number of methods of protein purification can be employed and these methods are known in the art and described extensively in the literature. For example, protein purification methods are dsescribed in Deutscher, 1990, Methods in Enzymology, 182(57):779; and Scopes, 1982, Methods in Enzymology 90: 479-91.

Often time, the selection of purification step(s) or methods depends, *e.g.*, on the nature of the production process, and the particular proteins that are produced.

6.6.4. Fermenting Microorganisms

The SSF methods of the disclosure employ a "fermenting microorganism" to generate a fermentation product (*e.g.*, ethanol) from the sugars produced in an attendant saccharification reaction and/or added to the system. Fermenting microorganisms capable of producing ethanol are sometimes referred to as ethanologens.

The term "fermenting microorganism," as used herein, refers to any microorganism suitable for use in a desired fermentation process. Suitable fermenting microorganisms according to the instant disclosure are able to ferment, *i.e.*, convert, sugars, such as, for example, glucose, xylose, arabinose, mannose, galactose, or oligosaccharides, directly or indirectly into the desired fermentation product.

Examples of suitable fermenting microorganisms include, without limitation, fungal organisms, such as yeast. Specifically, a suitable yeast can be selected from strains of the *Saccharomyces spp.*, and in particular, *Saccharomyces cerevisiae*. Various types of yeast are commercially available, among which, for example, ETHANOL RED™ yeast (available from Fermentis/Lesaffre, USA), FALI (available from Fleischmann's Yeast, USA), SUPERSTART™ and THERMOSACC® fresh yeast (available from Ethanol Technology, WI, USA), BIOFERM AFT and XR (available from NABC--North American Bioproducts Corporation, GA, USA), GERT STRAND (available from Gert Strand AB, Sweden), or FERMIOL (available from DSM Specialties) can be employed in performing the methods of the invention described herein.

In other aspects, the yeast is a *Saccharomyces distaticus* or a *Saccharomyces uvarum*. In yet other aspects, the yeast is a *Kluyveromyces*. Non-limiting examples of *Kluvermoyces* include *Kluyveromyces marxianus* or *Kluyveromyces fragilis*. In yet other aspects, the yeast is a *Candida*. Non-limiting examples of *Candida* include *Candida pseudotropicalis* and *Candida brassicae*. In yet other aspects, the yeast is a *Clavispora*. Non-limiting examples of *Clavispora* include *Clavispora lusitaniae* and *Clavispora*

opuntiae. In another aspect, the yeast is a *Pachysolen*, *e.g.*, a *Pachysolen tannophilus*. In another aspect, the yeast is a *Bretannomyces*, *e.g.*, a *Bretannomyces clausenii*. Yeast fermentation has been described in the literature. *See, e.g.,* Philippidis, 1996, Cellulose bioconversion technology, in Handbook on Bioethanol: Production and Utilization (Wyman, ed., Taylor & Francis, Washington, D.C., 179-212).

Bacteria that can efficiently ferment glucose to ethanol include, for example, *Zymomonas mobilis* and *Clostridium thermocellum* (*see, e.g.,* Philippidis, 1996, supra).

The cloning of heterologous genes in a *Saccharomyces cerevisiae* (*see, e.g.,* Chen and Ho, 1993, Appl. Biochem. Biotechnol. 39-40:135-147; Ho *et al.*, 1998, Appl. Environ. Microbiol. 64:1852-1859), or in a bacterium such as an *Escherichia coli* (*see, e.g.*, Beall *et al.*, 1991, Biotech. Bioeng. 38: 296-303), a *Klebsiella oxytoca* (*see, e.g.,* Ingram, *et al.*, 1998, Biotechnol. Bioeng. 58:204-214), or a *Zymomonas mobilis* (*see, e.g.*, Zhang *et al.*, 1995, Science 267:240-243; Deanda *et al.*, 1996, Appl. Environ. Microbiol. 62:4465-4470), has led to the construction of organisms capable of converting hexoses and pentoses to ethanol (cofermentation). Such microorganisms can advantageously be used in the methods of the present disclosure.

In certain embodiments, the fermenting microorganism is a *Zymomonas mobilis* with improved tolerance for acetate (*see, e.g.,* U.S. Patent Publication US 2009/0221078).

In certain embodiments, the fermenting microorganism is a *Zymomonas mobilis* with improved utilization of xylose (*see, e.g.,*, U.S. Patent Publication US 2009/0246846).

In certain embodiment, the fermenting microorganism is a *Zymomonas mobilis* with the ability to ferment pentoses into ethanol (*see, e.g.*, U.S. Patent Publication US 2003/0162271).

6.6.5. Fermentation Media

In some aspects, the SSF reactions or methods of the disclosure are performed in a fermentation medium or a complete fermentation medium. The term "fermentation medium," as used herein, refers to a medium before all of the components necessary for the SSF reaction to take place are present. A fermentation medium can thus be, for example, a medium resulting from a partial saccharification process. A fermentation medium can, in other embodiments, be a medium containing all the components necessary for the SSF reaction to take place. In that case, the fermentation medium is also termed "a complete fermentation medium." Moreover, a fermentation medium can, in yet other embodiments, be a medium wherein an SSF reaction is in progress or under way, and as such may contain certain products of saccharification.

A complete fermentation medium includes enzymes capable of hydrolyzing carbohydrate-based cellulosic or other substrates, a fermenting organism, and a carbohydrate-based cellulosic or other substrate (*e.g.*, as described in Section 6.7.2 below). Over the course of culturing the complete fermentation medium, fermentable sugars are formed through enzymatic hydrolysis, which are in turn metabolized by the fermenting organism to produce a fermentation product.

6.7. Simultaneous Saccharification and Fermentation Processes

In certain aspects, an SSF reaction of the present disclosure is performed at a temperature of between 25 °C and 50 °C. For example, the SSF reaction takes place at a temperature of 25 °C or above, 28 °C or above, 30 °C or above, 32 °C or above, 35 °C or above, or 38 °C or above. For example, the SSF reaction takes place at a temperature of 50 °C or below, 45 °C or below, 40 °C or below, 38 °C or below, 35 °C or below, or 30 °C or below. For example, the SSF reaction takes place in a temperature range of from 28 °C to 45 °C, such as from 30 °C to 40 °C, from 32 °C to 38 °C. In an exemplary embodiment, the SSF reaction is carried out at a temperature of from 32 °C to 35 °C. In another embodiment, the SSF reaction is carried out at a temperature of about 32 °C. The temperature at which the SSF reaction is carried out can, for example, be adjusted up or down during the reaction.

In SSF, the enzymatic hydrolysis of cellulose and the fermentation of glucose to ethanol are combined in one step (*see, e.g.,* Philippidis, 1996, Cellulose bioconversion technology, Handbook on Bioethanol: Production and Utilization, Wyman, ed., Taylor & Francis, Washington, D.C., pp. 179-212).

SSF processes are usually carried out as batch fermentation processes, wherein the fermentation is conducted from start to finish in a single tank. Alternatively, SSF processes can be carried out as continuous fermentation processes which are steady-state fermentation systems that operate without interruption, and wherein each stage of the fermentation occurs in a separate section of a given fermentation system, and flow rates are set to correspond to required residence times. In other words, the individual steps in a fermentation process of the disclosure can be performed batch-wise or continuously. Processes where all steps are performed batch-wise, or processes where all steps are performed continuously, or processes where one or more steps are performed batch-wise and one or more steps are performed continuously are contemplated herein.

In certain embodiments, a fed-batch SSF process may be desirable. A fed-batch process entails a batch phase and a feeding phase. The culture medium of the batch

phase and the culture medium added during the feeding phase are chemically defined, and the culture medium of the feeding phase is added, at least for a fraction of the feeding phase, at a feeding rate that follows a pre-defined exponential function, thereby maintaining the specific growth rate at a pre-defined value.

An SSF reaction of the present disclosure can suitably proceed for a period of 3 to 7 d. For example, an SSF reaction of the disclosure can proceed for up 3 d, 4 d, 5 d, 6 d, or 7 d.

The SSF fermentation processes of the disclosure include, without limitation, fermentation processes used to produce fermentation products including alcohols (*e.g.*, ethanol, methanol, butanol, 1,3-propanediol); organic acids (*e.g.*, citric acid, acetic acid, itaconic acid, lactic acid, gluconic acid, gluconate, lactic acid, succinic acid, 2,5 diketo-D-gluconic acid); ketones (*e.g.*, acetone); amino acids (*e.g.*, glutamic acid); gases (*e.g.*, H_2 and CO_2), and more complex compounds, including, for example, antibiotics (*e.g.*, penicillin and tetracycline); enzymes; vitamins (*e.g.*, riboflavin, B_{12}, β-carotene); hormones, and other compounds.

In certain aspects, the present disclosure provides a set of SSF conditions that are specifically suitable for use with a recombinant fermenting bacteria such as a Zymonmonas (*i.e.*, also termed "recombiant Zymomonas SSF conditions" herein). For example, these conditions include carrying out the SSF flask runs anaerobically under suitable recombinant *Zymomonas mobilis*, using pretreated substrates such as, for example, corn cob, bagasse, Kraft pulp substrate, and carrying out the reaction at about 33 ºC, pH 5.8, and about 10 wt.% to 25 wt.% solids loading, dependent upon the particular substrates and pretreatment. These conditions also include, for example, commencing the fermentation by the addition of 10% of a suitable *Zymomonas mobilis* strain, for example, strains ZW705 (recombinant) or ZW1 (wild-type) inoculum (5 g), into the reaction mixture without any additional nutrients.

In certain aspects, the present disclosure provides a set of SSF conditions that are specifically suitable for use with a fermenting microorganism that is a fungus, for example, a *S. cerevisiae* yeast (*i.e.*, also termed "yeast SSF conditions" herein). For example, these conditions include carrying out the reaction with a suitable yeast strain, for example the THERMOSACC® DRY yeast, at 38 ºC and pH 5.0, inoculation at 0.1 wt% without any additional nutrients, carrying out the SSF runs anaerobically by, for example CO_2 outgassing, using a reaction mixture comprising pretreated substrate, water, sulfuric acid, saccharification enzyme(s) and the yeast strain, as well as agitating

the reaction vessel at an appropriate speed, for example, at 100 RPM, for a suitable period of time, for example, 3 d.

6.7.1. Recovery of SSF Products

The fermentation product can be any substance that is produced by the fermenting organism. In a specific aspect, the substance is an alcohol. It will be understood that the term "alcohol" encompasses a substance that contains one or more hydroxyl moieties. In a specific aspect, the alcohol is arabinitol. In another aspect, the alcohol is butanol. In another aspect, the alcohol is ethanol. In another aspect, the alcohol is glycerol. In another aspect, the alcohol is methanol. In another aspect, the alcohol is 1,3-propanediol. In yet another aspect, the alcohol is sorbitol. In another more aspect, the alcohol is xylitol. *See, e.g.*, Gong *et al.*, 1999, Ethanol production from renewable resources, in Advances in Biochemical Engineering/Biotechnology, Scheper, ed., Springer-Verlag Berlin Heidelberg, Germany, 65: 207-241; Silveira and Jonas 2002, Appl. Microbiol. Biotechnol. 59: 400-408; Nigam, and Singh, 1995, Process Biochem. 30 (2): 117-124; Ezeji *et al.*, 2003, World J. Microbiol. Biotechnol. 19 (6): 595-603.

Distillation can be performed on the fermentation broth from the fermentation step to recover the fermentation products such as, for example, ethanol. The fermentation and distillation steps can be carried out simultaneously or separately/sequentially. In some aspects, after distillation, two products are recovered: an alcohol, such as, for example, ethanol, and a fermentation rest or residual product (whole stillage). The alcohol, being an azeotropic mixture with water, is further purified in the separation step by a standard process such as, for example, molecular sieving. For example, ethanol with a purity of up to about 96 vol.% can be obtained, which can be used as, for example, fuel ethanol, drinking ethanol, *i.e.*, potable neutral spirits, or industrial ethanol.

For other substances or fermentation products, any method known in the art can be used for recovery, including, but not limited to, chromatography (*e.g.*, ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (*e.g.*, preparative isoelectric focusing), differential solubility (*e.g.*, ammonium sulfate precipitation), SDS-PAGE, distillation, or extraction.

6.7.2. Sources of Carbohydrates or Feedstocks

Any suitable cellulosic substrates or raw materials can be used in practicing the SSF processes of the present disclosure. The substrate can be selected based on the desired fermentation product, *i.e.*, the substance to be obtained from the fermentation, and the process employed, as is well known in the art.

Examples of substrates suitable for use in the methods of present disclosure, include cellulose-containing materials, such as wood or plant residues, or low molecular sugars DP1-3 obtained from processed cellulosic materials that can then be metabolized by the fermenting microorganism, and/or which can be supplied by direct addition to the fermentation medium.

The biomass can include any composition comprising cellulose and/or hemicellulose (lignocellulosic biomass can also comprise lignin), *e.g.*, seeds, grains, tubers, plant waste or byproducts of food processing or industrial processing (*e.g.*, stalks), corn (including cobs, stover, and the like), grasses (*e.g.*, Indian grass, such as *Sorghastrum nutans*; or, switch grass, *e.g.*, *Panicum* species, such as *Panicum virgatum*), wood (including wood chips, processing waste), paper, pulp, recycled paper (*e.g.*, newspaper). Other biomass materials include, but are not limited to, potatoes, soybean (rapeseed), barley, rye, oats, wheat, beets or sugar cane bagasse.

6.8. Pretreatment of Biomass

Prior to an SSF reaction, biomass (*e.g.*, lignocellulosic material) is preferably subject to a pretreatment step in order to render xylan, hemicellulose, cellulose and/or lignin material more accessible to enzymes, and thus more amenable to saccharification and fermentation by the methods of the disclosure.

In one aspect, the pretreatment entails subjecting dried biomass material in a suitable vessel, such as for example, a reactor, to a catalyst comprised of a dilute solution of a strong acid and a metal salt; this can lower the activation energy, or the temperature, of cellulose hydrolysis to obtain higher sugar yields; *see, e.g.*, U.S. Patent Nos. 6,660,506; 6,423,145.

Another exemplary pretreatment method entails hydrolyzing biomass by subjecting the material to a first stage hydrolysis step in an aqueous medium at a temperature and a pressure level chosen to effectuate primarily depolymerization of hemicellulose without major depolymerization of cellulose to glucose. This step results in a slurry, in which the liquid aqueous phase contains dissolved monosaccharides resulting from depolymerization of hemicellulose, and the solid phase contains cellulose and lignin. A second hydrolysis step can involve conditions under which at least a major portion of the cellulose is depolymerized, resulting in a liquid aqueous phase containing dissolved/soluble depolymerization products of cellulose. *See, e.g.*, U.S. Patent No. 5,536,325.

Another exemplary method comprises processing a biomass material by one or more stages of dilute acid hydrolysis using about 0.4% to 2% of a strong acid; and

treating an unreacted solid lignocellulosic component of the acid-hydrolyzed biomass material by alkaline delignification. *See, e.g.*, U.S. Patent No. 6,409,841.

Another exemplary pretreatment method comprises prehydrolyzing biomass (*e.g.*, lignocellulosic materials) in a prehydrolysis vessel, for example, a reactor; adding an acidic liquid to the solid lignocellulosic material to make a mixture; heating the mixture to a suitable reaction temperature; maintaining the reaction temperature for a time period sufficient to fractionate the lignocellulosic material into a solubilized portion containing at least about 20% of the lignin from the lignocellulosic material and a solid fraction containing cellulose; removing a solubilized portion from the solid fraction while maintaining the mixture at or near the reaction temperature, wherein the cellulose in the solid fraction is rendered more amenable to enzymatic digestion; and recovering a solubilized portion. *See, e.g.*, U.S. Patent No. 5,705,369.

In another exemplary method, the pretreatment method uses hydrogen peroxide H_2O_2. *See, e.g.*, Gould, 1984, Biotech. Bioengr. 26:46-52.

In yet another exemplary method, pretreatment comprises contacting biomass with stoichiometric amounts of sodium hydroxide and ammonium hydroxide at very low biomass concentration. *See, e.g.*, Teixeira *et al.*, 1999, Appl. Biochem.Biotech. 77-79:19-34.

In another embodiment, pretreatment comprises contacting the 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, e.g.*, PCT patent application publication WO2004/081185.

In another exemplary method, the pretreatment uses ammonia. For example, the pretreatment method comprises subjecting the biomass to low ammonia concentration under conditions of high solids. *See, e.g.*, U.S. Patent Publication No. 20070031918, PCT patent application publication WO 2006/11901.

The invention is further illustrated by the following examples. The examples are provided for illustrative purposes only. They are not to be construed as limiting the scope or content of the invention in any way.

7. Example 1: analysis of Exp formation in SSF reactions

7.1. Materials & Methods

7.1.1. Substrates

Below is a list of substrates used in this work. The cellulose, xylan, and lignin compositions of the pretreated substrates are listed as well. Compositional analyses were performed using the standard assays detailed in the NREL protocols for Standard

Biomass Analytical Procedures (available at: http://www.nrel.gov/biomass/pdfs /42618.pdf):

Dilute ammonia pretreated corn cob. Corn cob was pretreated prior to enzymatic hydrolysis according to the methods and processing ranges described in, for example, US Patent Application Publications 2007-0031918-A1, US-2007-0031919-A1, US-2007-0031953-A1, US-2007-0037259-A1, and PCT patent application publication WO06/110901 A2 (unless otherwise noted). The composition of the substrate comprises: 34.8% cellulose, 29.2% xylan, 12.8% lignin.

Dilute sulfuric acid pretreated sugar cane bagasse. This substrate was produced and provided by NREL, as detailed in Schell *et al*, 2003, (App. Biochem. Biotechnol. Vol. 105-108, 69-85). The bagasse was pretreated at a solids concentration of 20% (w/w), temperature of 165 °C, 1.44% (w/w) acid and an approximate residence time of 8 min. The composition of the substrate comprises: 55.0% cellulose, 3.1% xylan, 31.2% lignin.

Mixed hardwood industrial unbleached pulp substrate. This substrate was produced using the Kraft process and oxygen delignification (Kappa Number = 13). (Research funded by l'Agence Nationale de la Recherche (ANR-05-BIOE-007) through l'Agence de l'Environnement et de la Maitrise de l'Energie (ADEME 0501C0099). The composition of the substrate comprises: 74.6% cellulose, 20.7% xylan, 2.6% lignin.

Softwood industrial unbleached pulp substrate. This substrate was produced using the Kraft process and oxygen delignification (Kappa Number = 14). (Research funded by l'Agence Nationale de la Recherche (ANR-05-BIOE-007) through l'Agence de l'Environnement et de la Maitrise de l'Energie (ADEME 0501C0099). The composition of the substrate comprises: 81.9% cellulose, 8.0% xylan, 1.9% lignin.

7.1.2. Enzymes

Below is a list of enzymes and enzyme mixtures used in this work.

Accellerase™ 1500 (Danisco U.S. Inc., Genencor) is a high β-glucosidase activity cellulase enzyme complex produced by a genetically modified *Trichoderma reesei*. It contains multiple enzyme activities a majority of which are exoglucanase, endoglucanase, β-glucosidase, and hemi-cellulase activities.

Multifect® Xylanase (Danisco U.S. Inc., Genencor), also produced by a *Trichoderma reesei*, is a hemicellulase enzyme complex designed to work as an

accessory product to supplement whole cellulase with xylanase activity, and to work synergistically to enhance various polysaccharide conversions in the lignocellulosic biomass processing industry. The predominant xylanase activity in Multifect® Xylanase is that of *T. reesei* Xyn2 (*see,* LaGrange *et al.*, 1996, Appl. Environ. Microbiol. 62:1036–1044).

Bxl1: is a β-xylosidase from *Trichoderma reesei*. The amino acid sequence of Bxl1 is provided herewith as SEQ ID NO:4. Bxl1 has been shown to have β-xylosidase activity in an assay using p-nitrophenyl-β-xylopyranoside, xylobiose or mixed, linear xylo-oligomers as substrates.

Xyn3: is a GH10 family xylanase from *Trichoderma reesei*. The amino acid sequence of Xyn3 is provided herewith as SEQ ID NO:18. Xyn3 was shown to have endoxylanase activity using birchwood azo-xylan (Megazyme, Wicklow, Ireland), and indirectly by its ability to increase xylose monomer production in the presence of xylobiosidase when Xyn3 in combination with xylobiosidase act on pretreated biomass or on isolated hemicellulose.

Fv3A: is a GH3 family enzyme from *Fusarium verticillioides*. The amino acid sequence of Fv3A is provided herewith as SEQ ID NO:6. Fv3A was shown to have β-xylosidase activity in an assay using p-nitrophenyl-β-xylopyranoside, xylobiose and mixed, linear xylo-oligomers (Figures 15 and 16) and branched arabinoxylan oligomers from hemicellulose as substrates.

Fv51A is a GH51 family enzyme from *Fusarium verticillioides*. The amino acid sequence of Fv51A is provided herein as SEQ ID NO:16. Fv51A was shown to have L-α-arabinofuranosidase activity in an assay using p-nitrophenyl- α-L-arabinofuranoside and by the release of arabinose from the set of oligomers released from hemicellulose by the action of endoxylanase.

Fv43D: is a GH43 family enzyme from *Fusarium verticillioides*. The amino acid sequence of Fv43D is provided herein as SEQ ID NO:2. Fv43D was shown to have β-xylosidase activity in an assay using p-nitrophenyl-β-xylopyranoside, xylobiose, or mixed, linear xylo-oligomers as substrates. (Figures 15 and 16).

Fv43B: is a GH43 family enzyme from *Fusarium verticillioides*. The amino acid sequence of Fv43B is provided herein as SEQ ID NO:12. Fv43E was shown to have β-xylosidase activity in an assay using p-nitrophenyl-β-xylopyranoside, xylobiose or mixed, linear xylo-oligomers as substrates.

Pf43A: is a GH43 family enzyme from *Penicillium funiculosum*. The amino acid sequence of Pf43A is provided herein as SEQ ID NO:8. Pf43A was shown to

have β-xylosidase activity in an assay using p-nitrophenyl-β-xylopyranoside, xylobiose or mixed, linear xylo-oligomers as substrates.

Fv43E: is a GH43 family enzyme from *Fusarium verticillioides*. The amino acid sequence of Fv43E is provided herein as SEQ ID NO:10. Fv43E was shown to have β-xylosidase activity in an assay using p-nitrophenyl-β-xylopyranoside, xylobiose or mixed, linear xylo-oligomers as substrates.

Af43A: is a GH43 family enzyme from *Aspergillus fumigatus*. The amino acid sequence of Af43A is provided herein as SEQ ID NO:14. Af43A was shown to have L-α-arabinofuranosidase activity in an assay using p-nitrophenyl- α-L-arabinofuranoside and by the release of arabinose from the set of oligomers released from hemicellulose by the action of endoxylanase.

XlnA: XlnA is a xylanase from *Aspergillus tubengensis*. The amino acid sequence of XlnA is provided herein as SEQ ID NO:20. The XlnA protein used in the present examples was unpurified, in the form of an enzyme prepration whose main constituent was XlnA.

Bgl1: Bgl1 is *T. reesei* β-glucosidase 1 (SEQ ID NO:26). The Bgl1 gene has been described, for example,in Barnett *et al.*, 1991, Bio-Technology 9(6):562-567.

7.1.3. Strains

Strain #229: A *Trichoderma reesei* strain, derived from RL-P37 (Sheir-Neiss and Montenecourt, 1984, Appl. Microbiol. Biotechnol. 20:46-53) through mutagenesis and selection for high cellulase production titer, was co-transformed with the β-glucosidase expression cassette (which comprised a *cbh1* promoter, *T. reesei* β-glucosidase1 gene, a *cbh1* terminator, and an *amdS* marker (*A. nidulans* acetamidase)), and the endoxylanase expression cassette (which comprised a *cbh1* promoter, *T. reesei xyn3,* and a *cbh1* terminator) using PEG mediated transformation (*see, e.g.,* Penttila *et al.*, 1987, Gene 61(2):155-64). Numerous transformants were isolated and examined for β-glucosidase and endoxylanase production. One transformant, referred to as *T. reesei* strain #229, was used in certain studies described herein.

Strain H3A: *T. reesei* strain #229 was co-transformed with the β-xylosidase Fv3A expression cassette (which comprised a *cbh1* promoter, an *fv3A* gene, a *cbh1* terminator, and an *alsR* marker (chlorimuron ethyl resistant mutant of the native *T. reesei* acetolactate synthase)), the β-xylosidase Fv43D expression cassette (which comprised an *egl1* promoter, an *fv43D* gene, a native *fv43D* terminator), and the Fv51A α-arabinofuranosidase expression cassette (which comprised an *egl1* promoter, an *fv51A*

gene, a native *fv51A* terminator) using electroporation (*see, e.g.,* PCT patent application publication WO2008/153712 A2). Transformants were selected on Vogels agar plates containing chlorimuron ethyl. Numerous transformants were isolated and examined for β-xylosidase and L-α-arabinofuranosidase production. *T. reesei* integrated expression strain H3A, which recombinantly expresses *T. reesei* β-glucosidase 1, *T. reesei* xyn3, Fv3A, Fv51A, and Fv43D, was isolated, and used in certain studies described herein.

7.1.4. Organisms and inoculum preparation

7.1.4.1. *Zymomonas mobilis*

Background: *Zymomonas mobilis* strain ZW1 is a wild type strain similar to strain ZM4 from American Type Culture Collection (ATCC 31821, Manassas, VA). Recombinant *Zymomonas mobilis* strain ZW705 was produced from strain ZW801-4 as summarized below. Cultures of *Z. mobilis* strain ZW801-4 were grown under conditions of stress as follows. ZW801-4 is a recombinant xylose-utilizing strain of *Z. mobilis* that was described in U.S. Patent Application Publication 2008/0286870. Strain ZW801-4 was derived from strain ZW800, which was, in turn, derived from strain ZW658, all as was described in U.S. Patent Application Publication 2008/0286870. ZW658 was constructed by integrating two operons, P_{gap}*xylAB* and P_{gap}*taltkt*, containing four xylose-utilizing genes encoding xylose isomerase, xylulokinase, transaldolase, and transketolase, into the genome of ZW1 (ATCC #31821) *via* sequential transposition events, and followed by adaptation steps conducted on selective media containing xylose. ZW658 was deposited as ATCC #PTA-7858. In ZW658, the gene encoding glucose-fructose oxidoreductase was insertionally-inactivated using host-mediated, double-crossover, homologous recombination and spectinomycin resistance as a selectable marker to create ZW800. The spectinomycin resistance marker, which was bounded by loxP sites, was removed by site specific recombination using Cre recombinase to create ZW801-4.

A continuous culture of ZW801-4 was grown in a 250 mL stirred, pH- and temperature-controlled fermentors (Sixfors; Bottmingen, Switzerland). The basal medium for fermentation was 5 g/L yeast extract, 15 mM ammonium phosphate, 1 g/L magnesium sulfate, 10 mM sorbitol, 50 g/L xylose, and 50 g/L glucose. Adaptation to growth in the presence of high concentrations of acetate and ammonia was achieved by gradually increasing the concentration of ammonium acetate in the above continuous culture media while maintaining an established growth rate as measured by the specific dilution rate over a period of 97 d to a concentration of 160 mM. Further increases in ammonium ion concentration were achieved by incremental additions of ammonium

phosphate, to a final total ammonium ion concentration of 210 mM by the end of 139 d of continuous culture. Strain ZW705 was isolated from the adapted population by plating, PCR amplification, and/or other conventional well-known methods.

7.1.4.2. Growth of seed cultures for SSF

Zymomonas mobilis strains ZW705 and ZW1 were maintained as 20% glycerol stocks, frozen at -80 °C. To make a seed culture, a 2 mL of this frozen stock was thawed and used to inoculate 45 mL of a medium containing 5 g/L of yeast extract, 4 g/L of potassium hydrogen phosphate, 1 g/L of magnesium sulfate, and 100 g/L of glucose at pH 5.8. The starting OD 600 nm was 0.4. The culture was grown at 32 °C in a capped 50 mL tube to an OD 600 nm of about 2.5, and it was then used to inoculate a final seed culture containing 200 g/L glucose, 4 g/L potassium hydrogen phosphate, 2 g/L magnesium sulfate, and 20 g/L yeast extract, at pH 5.8. That culture was grown at 32 °C in a pH-controlled and stirred fermenter to an OD 600 nm of about 10, and a remaining glucose concentration of about 80 g/L. A volume of this seed culture (having an OD 600 nm of about 10) equivalent to 10% of the SSF fermentation volume was used to start the SSF.

7.1.4.3. Yeast

The yeast ethanologen used was the THERMOSACC® DRY yeast (Ethanol Technology, Milwaukee, WI) which is only capable of fermenting C6 (*i.e.*, glucose) carbon surgars into ethanol (EtOH). The dry yeast was hydrated with sterile deionized water for 2-3 hrs prior to inoculation.

7.1.5. Fermentation using sugars

Zymomonas mobilis fermentation using synthetic sugars was carried out for 3 d in the 500 mL Sixfors bioreactor, using a batch and fed-batch process. The synthetic sugars consisted of glucose and xylose. For the batch process, the sugars were initially loaded at a concentration of about 80 g/L glucose and 70 g/L xylose. For the fed-batch process, a concentrated stock sugar solution was prepared first, which was then fed into the bioreactor using a syringe pump (PHD2000, Harvard Apparatus, Holliston, MA). The flow rate was set and controlled such that it gave a final equivalent sugar loading of about 80 g/L glucose and 70 g/L xylose at day 3. For the batch or the fed-batch processes, the *Zymomonas* inoculum was loaded into the reaction mixture in the beginning at 10 wt.% and the fermentations were carried out anaerobically at 33 ºC and pH 5.5.

7.1.6. Simultaneous saccharification and fermentation (SSF)

SSF flask runs were carried out anaerobically under suitable recombinant *Zymomonas mobilis* and yeast fermentation conditions. Unless otherwise stated, the recombinant *Zymomonas* SSF experiments using dilute ammonia pretreated corn cob, bagasse, or Kraft pulp substrate were carried out at 33 ºC, pH 5.8, and 25 wt.%, 20 wt.%, 10 wt.% solids loading, respectively. 25 wt.% solids (12.5 g dry weight), 20 wt.% solids (10.0 g dry weight), or 10 wt.% solids (5.0 g dry weight) of the respective substrates were each loaded first into a 125 mL Erlenmeyer flask, followed by the addition of deionized water pre-mixed with the required amount of 6N sulfuric acid, in order to titrate the substrate pH to 5.8. The cellulase and hemicellulase enzymes, described above, were added in an amount based on mg cellulase protein/g cellulose and mg hemicellulase protein/g xylan in the biomass substrate, respectively. Fermentation was initiated by the addition of 10 wt.% *Zymomonas mobilis* strains ZW705 (recombinant) or ZW1 (wild-type) inoculum (5 g) into the reaction mixture without any additional nutrients. For the THERMOSACC® DRY yeast SSF, the reactions were carried out at 38 ºC and pH 5.0. The level of yeast inoculation was at 0.1% w/w without any additional nutrients. The anaerobic environment and CO_2 outgassing were maintained by a 23 Gauge needle protruding from a rubber stopper that was used to cap the flask. At the start of fermentation, all SSF runs had an initial 50 g total reaction weight in a flask and the reaction mixture consisted of pretreated substrate, water, sulfuric acid, enzyme, and either *Zymomonas* or yeast cells. The flasks were agitated inside a shaker incubator (New Brunswick Scientific, Innova 44, Edison, New Jersey) at 100 RPM for 3 d.

7.1.7. Separate hydrolysis and fermentation (SHF)

The SHF runs involved a saccharification stage that was followed by a fermentation stage. The saccharification conditions were based on the NREL Laboratory Analytical Procedure (*see,* Selig *et al.*, 2008, Enzymatic Saccharification of Lignocellulosic Biomass Laboratory Analytical Procedure (LAP), Technical Report NREL/TP-510-42629) except for certain modifications of the enzyme types and/or levels, cellulose loading, and pH. 25 wt.% solids (12.5 g dry weight) of dilute ammonia pretreated corn cob or 10 wt.% solids (5.0 g dry weight) of the mixed hardwood pulp was loaded into a 125 mL Erlenmeyer flask. This was then followed by the addition of deionized water pre-mixed with the required amount of 6N sulfuric acid so as to titrate the substrate pH to 5.3. The saccharification step was commenced by the addition of cellulase and hemicellulase enzymes in an amount based the total mg protein/g cellulose

and the total mg protein/g xylan in the biomass substrate, respectively. For all *Zymomonas* SHF runs, a temperature of 50 °C and a pH of 5.30 were used for the saccharification at a duration of 3 d, followed by the fermentation step using similar conditions as those described in the SSF process (above). Sodium hydroxide was used to raise the pH from 5.3 to 5.8. This was followed by the addition of 10 wt.% of *Zymomonas* inoculum to commence fermentation.

7.1.8. Fed-batch SSF

The fed-batch SSF studies were carried out under similar conditions as described in the SSF process (above) except that the dilute ammonia pretreated corn cob substrate (required to achieve a final 25 wt.% solids) was divided evenly into 8 batches, and fed batch-wise into the bioreactor, with the final batch loaded at the 30th hr.

7.1.9. HPLC analysis and EXP quantification

Fermentation samples were taken at timed intervals and analyzed for ethanol, residual sugars, ethyl-β-xylopyranoside (EXP), and other metabolic products, such as, for example, acetic acid and glycerol, using a Waters HPLC system (Alliance system, Waters Corp., Milford, MA). The HPLC column was purchased from BioRad (Aminex HPX-87H, BioRad Inc., Hercules, CA). The EXP quantification under refractive index detection followed a set of procedures that were similar to those described in by Zhang *et al.* (in 2009, Enzyme and Microbial Technology 44:196–202). Another metabolic co-product, succinic acid, was found to co-elute with the EXP and to have potentially inflated the EXP quantification. It was determined that, for the yeast and *Zymomonas* fermentations, the concentrations of succinic acid generated at the conditions tested were no more than 1-2 g/L, as measured using both an UV detector at 220 nm and an enzymatic assay kit (K-SUCC, Megazyme, Co. Wicklow, Ireland).

7.2. Results

7.2.1. EXP formation and identification

7.2.1.1. EXP formation in SSF with recombinant *Zymomonas mobilis*

Formation of a byproduct was observed under the SSF fermentation conditions described above, using a recombinant *Zymomonas mobilis* strain that is capable of co-fermenting glucose and xylose into ethanol. The substrate used in this study was dilute ammonia pretreated corn cob, which has a high xylan content, treated with commercial cellulase/hemi-cellulase enzyme preparations (Accellerase™ 1500 at 20 mg/g cellulose) and Multifect® Xylanase (5 mg/g xylan), both derived from *Trichoderma reesei* (*T. reesei*). Under the conditions tested, the highest amount of formation of the byproduct

was found under fed-batch SSF, and the second highest in SSF (Table 1). SHF using dilute ammonia pretreated corn cob, batch, and fed-batch fermentation processes using sugars (80 g/L glucose + 70 g/L xylose) formed little of this byproduct.

The byproduct generated under the SSF conditions, using dilute ammonia pretreated corn cob and the *Zymomonas mobilis* strain as described above had an elution time on an HPLC HPX-87H column of approximately 11.75 mins (at a flow rate of 0.6 mL/min), close to that of succinic acid. Figure 1 panel 2 shows a peak eluting at 11.611 min, following incubation of xylose and Multifect® Xylanase with ethanol. That particular peak was absent when no alcohol was present in the incubation. It was observed that the position of the byproduct generated in the presence of alcohol during incubation shifted to an extent and in a direction that tracked the alcohol added. The elution times of the byproduct produced after incubation with methanol (MeOH) at19.46 min, and the byproduct produced after incubartion with n-PrOH at 27.65 min were shorter and longer, respectively, when compared to the byproduct produced after incubation with ethanol (EtOH), at 21.99 min. The elution times of the products produced by incubation of xylose with Multifect® Xylanase and 0.72 M alcohol shift, relative to the ethanol-induced product (at 11.61 min), to shorter elution times with MeOH (at 11.03 min) and longer elution times with n-PrOH (at 13.60 min). These results were consistent with the conclusion that the mobile peaks in question were xylose-alcohol adducts (alkyl-xylopyranosides) formed by reverse hydrolysis from xylose and alcohol, with the components eluting at 11.03, 11.61 and 13.60 corresponding to the methyl-, ethyl- and n-propyl-xylopyranosides, respectively.

7.2.1.2. Time Course

The time course for the appearance of the methyl-, ethyl- and n-propyl xylopyranosides is shown in Figure 2. The relative amounts of the products formed after 100 hrs were as follows: methyl->ethyl->n-propyl-xyloside (Figure 2), consistent with the order reported by Drouet *et al.* (in 1994, Biotech. & Bioeng. 43:1075-1080) for the formation of the alkyl-β-D-xylopyranosides by reverse hydrolysis in the presence of MeOH, EtOH, and n-PrOH. The coincidence of the elution times of the ethyl xylopyranoside (EXP) withthose observed under the SSF conditions, the dependencies of the elution time on the presence and nature of the alcohol, and the relative reactivities of these alcohols, all suggested that the EXP was the byproduct generated under the SSF conditions.

7.2.1.3. Identification of ethyl-β-glycosides in an EXP sample by ^{1}H NMR analysis

Preparation of ethyl b-D-xylopyranoside standard: A sample of ethyl xylosides was prepared by dissolving 50 mg of D-xylose in 3 mL of ethanol and heating the resulting solution in the presence of Amberlyst 15 H^+ resin (100 mg) at 70 °C for 3 hrs. The resin was filtered and the ethanol solvent removed under reduced pressure, yielding a colorless oil. Analysis by ^{1}H NMR revealed a mixture of ethyl xylosides, with a-D-xylopyranoside predominating, along with amounts of ethyl β-xylopyranoside and ethyl α/β-xylofuranoside.

^{1}H NMR of EXP Sample: A lyophilized EXP sample purified and fractionated (using HPLC) from a dilute ammonia pretreated corn cob SSF fermentation broth sample was reconstituted into 750 µL of D_2O and transferred to a PP-528 glass NMR tube. A proton NMR spectrum was acquired on a Varian 500 MHz VNMRS NMR system using a basic s2pul pulse sequence for over 8 pulses. The ^{1}H NMR spectrum (500 MHz, D_2O), referenced to the HOD signal at *d* 4.80, indicated the presence of at least 4 distinct ethyl glycosides, based on the appearance of several triplet signals in the *d* 1.20-1.26 region. The spectrum also contained 5 distinct doublets at *d* 4.43, 4.49, 4.53, and 4.66 with coupling constants of 7.8 Hz indicative of *b*-glycosidic linkages. The signal at d 4.43 was the most intense and matches the chemical shift reported in the literature (Drouet *et al.*, 1994, Biotech. & Bioeng. 43:1075-1080), Zhang *et al.* (2009, Enzyme and Microbial Technology 44:196–202) for ethyl-β-D-xylopyranoside. Overall the signals observed in the ^{1}H NMR spectrum (Figure 3) indicated the presence of the β-anomer, β-D-ethylxylopyranoside, along with 3-4 additional ethyl β-glycosides.

7.2.2. EXP formation with yeast and wild-type *Zymomonas mobilis*

Co-fermentation of C5/C6 sugars using the ethanologen *Zymomonas mobilis* under SSF and fed-batch SSF conditions with a dilute ammonia pretreated corn cob substrate resulted in high levels of EXP. A new experiment was designed to determine whether EXP is formed when SSF is performed using wild-type yeast and *Zymomonas mobilis*, which are organisms that are only capable of fermentingC6 sugars. The results (Figure 4) indicated that the EXP was produced at high levels in the SSF reactions wherein the yeast (5.5 g/L) or wild-type *Zymomonas mobilis* (9.1 g/L) did the fermenting. Thus, the EXP formation is not specific to the recombinant *Zymomonas mobilis*, as it was also detected during SSF by a *Saccharomyces cerevisiae* yeast and a wild-type *Zymomonas mobilis*.

7.2.3. EXP formation from other sources of biomass

Additional experiments were designed to determine if EXP could be formed using substrates other than the dilute ammonia pretreated corn cob. From the results shown in Table 2, the formation of EXP appears to be linked to the fermentation of any high xylan-containing substrates under the SSF conditions with the *T. reesei* cellulase/hemicellulase preparations. For example, as much as 16.1 g/L EXP was formed under the recombinant *Zymomonas* SSF conditions using a mixed hardwood industrial Kraft pulp with high xylan content. On the other hand, the lower xylan-containing substrates, such as the pretreated bagasse and the softwood pulp, produced lower amounts of EXP (0.91 and 4.75 g/L, respectively) under similar conditions. Furthermore, it was again demonstrated that, only under the SSF processes as described herein would large amounts of EXP be generated. Formation of EXP was much decreased under the SHF process conditions, (3.52 g/L), even when a higher xylan-containing substrate was used.

7.2.4. EXP formation with different permutations of enzymes used in SSF

7.2.4.1. EXP formation with BxI1

According to Drouet *et al.* (in 1994, Biotech. & Bioeng. 43:1075-1080), EXP formation was catalyzed by *T. reesei* BxI1 *via* both a transxylosylation and a reverse hydrolysis reaction. Figure 5 shows that the EXP formation under the yeast SSF conditions described above was enhanced in the presence of *T. reesei* BxI1. With the addition of 5 mg/g BxI1, the EXP production was increased from 5.5 g/L to 18.8 g/L.

EXP formation under recombinant *Zymomonas mobilis* SSF conditions with 25 wt.% solids of dilute ammonia pretreated corn cob substrate was also enhanced in the presence of *T. reesei* BxI1, as shown in Figure 6. EXP plateaued at a concentration of 25.6 g/L, with the addition of 6 mg/g *T. reesei* BxI1. A cellulase/hemi-cellulase enzyme complex from *T. reesei* integrated strain #229 (which overexpresses xylanase, *T. reesei* Xyn3) was used in this experiment in place of Accellerase™ 1500 + Multifect® Xylanase. The ability of *T. reesei* BxI1 to catalyze EXP formation is surprisingly strong, because only 1 mg/g of it was added but an over 2-fold increase of EXP as compared to the control sample (enzyme complex from *T. reesei* integrated strain #229 alone) was clearly observed.

The effect of *Fusarium verticillioides* hemi-cellulase addition on EXP formation was investigated. Accellerase™ 1500 + Multifect® Xylanase, Accellerase™ 1500 + XlnA, and enzyme complex from the integrated *T. reesei* strain #229, with the addition of

BxI1, and *Fusarium verticillioides* hemi-cellulases (Fv3A, and Fv51A L-α-arabinofuranosidase were used. This study was conducted under recombinant *Zymomonas* SSF conditions. The substrate used was a 25 wt.% solids dilute ammonia pre-treated corn cob. Results were shown in Figure 7, which indicated that the addition of XlnA and Accellerase™ 1500 produced more EXP (31.5 vs. 19.5) than the addition of Multifect® Xylanase and Accellerase™ 1500. The enzyme complex from the integrated *T. reesei* strain #229 alone produced the least amount of EXP (9.1 g/L) among all three enzyme configurations. This is possibly due to the fact that BxI1 represented a smaller fraction of the total amount of proteins in the enzyme complex in the *T. reesei* strain #229 enzyme complex, which also produced a relatively large amount of accumulated xylobiose , as compared to other two enzyme configurations. For Accellerase™ 1500 + Multifect® Xylanase and Accellerase™ 1500 + XlnA, the addition of Fv3A or Fv51A did not result in substantial increases in EXP formation. With the enzyme complex from integrated *T. reesei* strain #229, however, the addition of Fv3A alone gave a 2-fold increase in EXP formation while the addition of Fv51A increased EXP formation by 1.4 fold, as compared to a 3-fold increase in EXP when *T. reesei* BxI1 was added to the enzyme complex produced from the integrated *T. reesei* strain #229.

7.2.4.2. EXP formation using purified enzymes

The results discussed above indicated that the EXP formation was strongly affected and effectively catalyzed by *T. reesei* BxI1 and that *T. reesei* BxI1 is remarkably and particularly effective at making EXP under the SSF conditions using certain high xylan-containing biomass substrates. However, all of the enzymes tested so far (and discussed above) were not purified and might contain background enzyme activities similar to that of the *T. reesei* BxI1, capable of making EXP. To investigate the effect of background enzyme activities, the effect of purified enzymes on EXP formation was also studied. A cellulase mixture of purified *T. reesei* cellobiohydrases, CBH1 and CBH2; *T. reesei* endoglucanase, EG2; and *T. reesei* β-glucosidase, BgI1, was used as a substitute for Accellerase™ 1500, while purified *T. reesei* Xyn3 was used as a substitute for Multifect™ Xylanase. It is clear from the results depicted in Figure 8 that the cellulase alone does not produce EXP. The addition of unpurified *T. reesei* Xyn3 produced large amounts of EXP, for example, about 13.1 g/L. This large increase is potentially due to a larger background *T. reesei* BxI1 that exists within the unpurified *T. reesei* Xyn3 sample.

7.2.4.3. EXP formation using GH43 class BxI enzymes

BxI1, which is a GH3 family hydrolases, has being shown to be active and effective at catalyzing the formation of EXP under the SSF conditions (above). A

question remains as to whether other GH family β-xylosidases can also catalyze the formation of EXP under similar conditions. A number of β-xylosidases from the GH43 family, including Fv43B, Pf43A, Fv43E, and Af43A, were tested under the recombinant *Zymomonas* SSF conditions using a 25 wt.% solids dilute ammonia pretreated corn cob substrate. Pf43A, Fv43E, and Af43A were found to increase EXP formation slightly, as compared to the control sample made from a protein preparation of strain #229. On the other hand, Fv43B gave a greater increase of EXP formation by 1.5-fold (Figure 9). Because all of these GH43 family enzymes were expressed in *T. reesei,* and because they were all unpurified protein preparations in this study, it could be postulated that the increase in EXP formation may be attributed to the presence of native Bxl1 in the protein preparations.

8. EXAMPLE 2: REDUCTION IN EXP BY HEMICELLULASES

The following examples show that EXP is reduced by addition of certain hemicellulases to SSF reactions.

8.1. Reduction in EXP by Fv43D in SSF by *Zymomonas*

EXP formation typically consumes both xylose (directly, or from xylobiose or other xylo-oligomers) and ethanol on an equal molar basis. This consumption mechanism directly results in a substantial decrease in yield for ethanol, because many microorganisms, including *Zymomonas,* are incapable of degrading and fermenting, or otherwise utilizing EXP. One (1) g of EXP is calculated to be equivalent to a 0.688 g loss of ethanol that would have been produced (assuming xylose is fermented into ethanol at a rate of 0.51 g/g xylose). Thus preventing the formation of EXP can lead to an attendant increase in ethanol yield. It was found, as shown in Figure 10, that the addition of the *Fusarium verticillioides* β-xylosidase, Fv43D, at only 1 mg/g xylan, greatly reduced (~4-fold) the amount of EXP formed under the recombinant *Zymomonas* SSF conditions.

8.2 Reduction in EXP by Fv43D in SSF by yeast

Similar to the results obtained from the C5/C6-fermenting recombinant *Zymomonas*, the supplementation of Accellerase™ 1500+Multifect® Xylanase with Fv43D in an SSF reaction using yeast as a C6-fermenting microorganism also resulted in a reduction in EXP formation. Specifically the addition of Fv43D at 1 mg/g xylan resulted in a 2 to 3-fold reduction in EXP (Figure 11).

8.3 Reduction in EXP by Fv43D in SSF with Multifect xylanase, with XlnA and with enzyme complex from *T. reesei* integrated strain #229

The reduction in EXP by the addition of Fv43D to three enzyme configurations, Accellerase™ 1500+Multifect® Xylanase, Accellerase™ 1500+*A. niger* xylanase, and the enzyme complex produced from the integrated *T. reesei* strain #229, under the recombinant *Zymomonas* SSF conditions was investigated. The substrate used was a 25 wt.% solids dilute ammonia pretreated corn cob.

Figure 12 shows that the addition of only 1 mg/g xylan of Fv43D to Accellerase™ 1500+Multifect® Xylanase, and to Accellerase™ 1500+XlnA, resulted in an over 4-fold reduction in EXP formation, and the same addition to the enzyme complex produced from the integrated *T. reesei* strain #229 resulted in a 1.36-fold reduction of EXP formation. At the same time, a corresponding increase in ethanol was observed with all three enzyme configurations, confirming the benefit of Fv43D in terms of reducing EXP formation and preventing ethanol yield loss.

8.4 Reduction in EXP formation in SSF with purified enzymes

To insure that the reduction of EXP from Fv43D was not due to background enzyme activities, purified enzymes including a purified Fv43D were used to test for reductions of EXP formation. A *T. reesei* cellulase mixture of purified CBH1, CBH2, EG2 and Bgl1 was used to substitute for Accellerase™ 1500, and a purified *T. reesei* Xyn3 was used as a substitute for Multifect® Xylanase. The results as depicted in Figure 13 showed a slight reduction in EXP formation and an attendant slight increase in ethanol titer, when purified and unpurified Fv43D were added. The relatively small reduction in EXP formation is likely due to the lack of background *T. reesei* Bxl1 in the purified enzymes (e.g., the cellulases and Xyn3). It is also noted that only a small amount of EXP (5.7 g/L) was formed by the control sample, due to the action of *T. reesei* Xyn3. Further addition of Fv43D thus does not substantially reduce EXP formation. To investigate this further, another study with results shown in Figure 14A was performed to investigate the effect of EXP reduction by Fv43D addition in the presence of a large amount of Bxl1.

8.5 EXP reduction dose response from the addition of Fv43D to SSF

EXP formation in the presence of Bxl1 was substantial where more than 20 g/L of EXP was consistently detected, although the exact amount varies by the SSF conditions used. A dose response study was performed to assess reduction of EXP formation in relation to the amount of Fv43D added. Figure 14A shows the results for EXP reduction from the addition of increasing amounts of Fv43D to the enzyme complex produced from

the *T. reesei* integrated strain #229 and *T. reesei* BxI1. Similar to the results shown previously, 1 mg/g of Fv43D was found to be effective at reducing EXP formation by nearly 3-fold and at the same time resulted in an increase in ethanol titer. However, addition of increasing amounts of Fv43D at 3, 6, and 9 mg/g did not greatly reduce EXP formation, although a significant increase in ethanol titer was observed. It appeared that even with addition of a large amount of Fv43D (at 9 mg/g), the EXP concentration could not be reduced to below 6.6 g/L, a level at which the amount of EXP may have reached equilibrium for the specific conditions that were used for testing. Under these particular experimental conditions, further additions of Fv43D beyond, for example, 1 mg/g xylan, no longer had an effect on reducing EXP formation.

These observations can potentially be explained by the mechanism of EXP formation and reduction under SSF conditions.

The proposed mechanism, which is based on the observations made from this work and is consistent with the data reported in the literature (*see, e.g.,* Drouet *et al.*, 1994, Biotech. & Bioeng. 43:1075-1080 and Zhang *et al.*, 2009, Enzyme and Microbial Technology 44:196–202), is described in detail in Example 3 below.

8.6 EXP reduction by the addition of Fo43A and Gz43A to SSF

Fo43A and Gz43A, which are inverting GH43 family β-xylosidases, were also tested for efficacy in reducing EXP formation under the SSF conditions described herein. Based on the results shown in Figure 14B, these two enzymes were each able to reduce EXP formation by nearly 4-fold, when they are added to a enzyme blend comprising a protein complex produced from the *T. reesei* integrated strain #229, *T. reesei* BxI1, a dilute ammonia pretreated corn cob substrate, and a recombinant *Zymomonas*, and at 3 mg/g xylan. Similar to what was observed with the addition of Fv43D (above), a reduction in EXP formation and a corresponding increase (by about 10 g/L) in ethanol titer was observed.

9. EXAMPLE 3: MECHANISM OF EXP FORMATION

Observations that the yield of EXP was approximately four times higher under the SSF conditions (Figures 10 and 11) in the presence of *Fusarium verticillioides* β-xylosidase Fv3A as compared to the yield of EXP in the presence of *Fusarium verticillioides* β-xylosidase Fv43D led to an effort to understand the mechanism of EXP formation.

Potentially there are two possible routes for the formation of EXP: (1) by transglycosylation; or (2) by reverse hydrolysis. Each of these routes are depicted as follows:

1) *Transglycosylation*: X-O-X + EtOH → X-O-Et + X

2) *Reverse Hydrolysis*: X + EtOH ↔ X-O-Et + H_2O

According to Drouet *et al.* (in 1994, Biotech. & Bioeng. 43:1075-1080), the transglycosylation mechanism is the more rapid of the two. The fact that only the β anomer (see Figure 29) is formed in a transglycosylation reaction implies that the reaction is catalyzed by enzymes that operate with retention of anomeric configuration on the substrates, wherein the sugars are linked by β-glycosyl linkages. On the other hand, formation of the β anomer of EXP by reverse hydrolysis would suggest that the starting substrate is either α- or β-xylose, depending on whether EXP is formed, respectively, by inversion or retention of anomeric configuration.

The β-xylosidases, Fv3A (a member of the GH3 family) and Fv43D (a member of the GH43 family), operate with retention and inversion of configuration, respectively. To distinguish and identify which of the two potential mechanisms described above is the true mechanism, an experiment was performed in which Multifect® Xylanase (which is an enzyme preparation characterized by the retaining activity of *T. reesei* Xyn2) and purified Fv43D and Fv3A were incubated at 46 ºC with xylobiose and ethanol. The kinetics of EXP formation in these samples was examined by HPLC.

As shown in Figure 15 and Table 3, after a short period of time Multifect® Xylanase with Fv3A produced a substantial amount of EXP, which appeared to correspond with xylose formation (e.g., having a constant ratio of EXP formation to xylose formation with time). When Fv43D was used with Multifect® Xylanase, on the other hand, EXP appeared with a substantial lag relative to xylose formation. It can be postulated that Multifect® Xylanase and Fv3A produced EXP rapidly through transglycosylation. A xylose-enzyme ester adduct would have been formed first with C1 anomeric inversion, followed by a second inversion by ethanol doing an S_N2 displacement, at the C1 carbon of the attached xylose, of the active site carboxyl group. The double inversion gives retention of configuration in the product. On the other hand, Fv43D did not appear to catalyze a transglycosylation reaction; instead it produced EXP much more slowly though reverse hydrolysis. As Fv43D operates with inversion of anomeric configuration, EXP was likely formed from α-D-xylose, which in turn was formed from xylobiose by Fv43D, but also by inversion at C1 of β-D-xylose in an aqueous solution. Figure 15 indicates that EXP can also be formed by reverse hydrolysis from xylose, catalyzed by Multifect® Xylanase. Because the β-xylosidase

mechanism in this case was retaining, the likely substrate of this reaction was β-D-xylose.

9.1 Equilibrium equation for EXP formation

Under the SSF conditions, Multifect® Xylanase and Fv3A would encounter not only xylobiose but larger xylose oligomers as well. Hydrolysis of such oligomers could result in an increased formation of EXP, if upon scission, one of the oligomer products is transglycosylated to produce an ethyl glycosyl adduct. Further scission, by transglycosylation, could then form an additional ethyl adduct such that the ratio of EXP formation to xylose formation would be greater than that observed when only xylobiose is transglycosylated. For example:

3) X4 + EtOH → X2-Et + X2

4) X2-Et + EtOH → 2EXP

5) X2 + EtOH → EXP + X

It is noted that if only xylobiose is transglycosylated, the only EXP formation is from the reaction of line 5).

These possibilities were examined by comparing hydrolysis of xylose oligomers (average molecular weight = 539) with that of xylobiose using each of the same three enzymes: Fv3A, Fv43D, and Multifect® Xylanase. The results are depicted in Figure 16.

Plotted in Figure 16 (left panel) and Figure 16 (right panel) are the ratios of EXP to Xylose RI peak areas generated following incubation of xylobiose (20 mg/mL) (Figure 16 (left)) or xylose oligomer (20 mg/mL) (Figure 16 (right)) in the presence of 0.9 M EtOH at 46 ºC in the presence of Multifect® Xylanase (560 µg/mL), Fv43D (36 µg/mL), or Fv3A (54 µg/mL), at concentrations that would give rates of xylose formation that were within a factor of two of each other. The EXP formation rates were nearly the same for xylobiose and xylose oligomers in the presence of Multifect® Xylanase and Fv43D. However, in the case of Fv3A, the EXP yield was about 1/3 higher in the presence of the xylose oligomers as compared to in the presence of xylobiose. As outlined above, this observation suggests that, at least for Fv3A, transglycosylation can occur with ethanol upon scission of oligomers larger than a xylobiose. Such transglycosylation of ≥dp3 oligomers does not appear to occur as prevalently in the cases of Multifect® Xylanase, although the EXP yield did appear to be higher than that was observed for Fv3A in both the xylobiose and xylose oligomer cases. This higher yield is likely an indication of a higher rate of transglycosylation with ethanol for Multifect® Xylanase than for Fv3A. The

fact that for both Multifect® Xylanase and Fv3A, EXP is formed for every 2.5 to 3.5 scission reactions is remarkable, considering that the molarity of ethanol is 0.9 M as compared to 55 M for H_2O. The selectivity for ethanol likely indicates a greater affinity for ethanol over water at or near the active site of the particular enzyme, with that affinity being higher for Multifect® Xylanase relative to that for Fv3A. The ratio of EXP to xylose is the same for Fv43D with xylobiose or xylose oligomer as substrate, which can be explained by the fact that, in this the case of Fv43D, EXP is formed by reverse hydrolysis from xylose, which was the product of hydrolysis from both xylobiose and xylose oligomer substrates.

In the case of Multifect® Xylanase and Fv3A, the maximum yield of EXP is attained after 3.5 hrs of incubation. The EXP yield at this point of incubation using Multifect® Xylanase was about 7 to 8 times that of the EXP yield obtained using Fv43D. After 3.5 hours the yield of EXP decreased in the Multifect® Xylanase reaction and the Fv3A reaction with a $t_{1/2}$ of 40-60 h. The reaction involving Fv3A was likely the enzyme-catalyzed hydrolysis of EXP from the high value formed in the transglycosylation reaction toward the equilibrium value of EXP/xylose formed upon reverse hydrolysis. In SSF reactions, the speediness of EXP formation and the sluggishness of the subsequent EXP hydrolysis in the presence of β-xylosidases having retaining activities suggests that the concentration of EXP is likely to remain substantially higher throughout the duration of an SSF reaction when a retaining β-xylosidase is present than when an inverting β-xylosidase is present. This further suggests that replacing retaining β-xylosidases with inverting β-xylosidases in SSF reactions would be beneficial for product yield. It is noted that all of the above mentioned reactions are enzyme catalyzed. Control samples/reactions, in the absence of enzyme, showed no formation of EXP upon incubation of xylobiose, of xylose oligomer, or of xylose in the presence of 0.9M ethanol at 46°C.

9.2 Calculation of equilibrium constant

It has been suggested by Drouet *et al.* (in 1994, Biotech. & Bioeng. 43:1075-1080) that the extent of formation of the alkyl-xylopyranoside (AXP), for example, EXP, upon reverse hydrolysis can be determined by the equilibrium established between the alcohol, xylose and the alkyl-xylopyranoside product.

The dissociation constant for ethyl-xylopyranoside (EXP) would be:

$$Kd = \frac{[xylose][EtOH]}{[EXP]}$$

Using this formula, a Kd of 27 M can be calculated from the data of Figure 2 as the equilibrium constant when Multifect® Xylanase was present in the SSF reaction. Also using this formula, a Kd of about 9.4 M can be calculated from the data of Figure 16 as the equilibrium constant when Fv43D was present in the SSF reaction. From Drouet *et al.* (1994, Biotech. Bioeng. 43:1075-1080), used *T. reesei* β-xylosidase and calculated a Kd of 40 M, although it did not appear that the EXP concentration had reached the equilibrium constant even after 160 hrs of incubation. The equilibrium constant calculations above were all made from SSF reactions started with xylose and ethanol, with the exception of the results depicted in Figure 16 of the disclosure, where the xylobiose and xylose oligomers were completely hydrolyzed by Fv43D in 20 hrs but the SSF reaction was allowed to proceed for a total of 145 hrs. Thus it is postulated that the Kd falls in the range of between about 10 and about 40M and is probably closer to the lower value, because the inverting β-xylosidase Fv43D produces a significant amount of EXP by the reverse reaction at equilibrium.

10. EXAMPLE 4: DELETION OF *BXL1* GENE FROM *T. REESEI*

10.1 Construction of the *bxl1* deletion cassette

To construct the *bxl1* deletion cassette, 5' and 3' flanking sequences of the *bxl1* gene (Margolles-Clark *et al.*, 1996, App. Environ. Microbiol. 62(10):3840-6) from *Trichoderma reesei* genomic-DNA were amplified by PCR with primer pairs MH375/MH376 and MH377/MH378 respectively (shown in Table 4), using PfuUltra II Fusion HS DNA Polymerase (Stratagene). The 3'-flanking sequence contained part of the Bxl1 coding sequence to avoid the nearby *bgl1* gene. Primer MH376 was phosphorylated at the 5'-end. One (1) µL of T4 DNA Ligase at a concentration of 5 U/µL (Roche Applied Bioscience), 1 µL of 10X ligation buffer (Roche) and approximately 20 ng of the PCR fragments were incubated for 10 mins at room temperature in a total volume of 10 µL. The ligation reaction mixture was used as a template for a PCR reaction with primer pair MH379/MH380 and PfuUltra II Fusion HS DNA Polymerase (Stratagene).

The resulting 4.0 kb fragment was cloned into pCR-Blunt II-TOPO according to manufacturer's specifications (Invitrogen). The plasmid was transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen). A colony, which contained the 4.0 kb *bxl1* 5'+3' PCR fusion product cloned into the TOPO vector, was isolated. The plasmid was extracted by QiaPrep plasmid purification (Qiagen) and its sequence confirmed (Sequetech, Mountain View, CA). The resulting plasmid was digested with *Acl*

and *Asc*I (NEB) to allow for subsequent cloning with the fragment containing the hygromycin-resistance gene.

The hygromycin resistance gene was amplified with primers MH292/MH290 from a vector containing the *Aspergillus nidulans oliC* promoter, the hygromycin resistance gene (*E. coli,* hygromycin phosphotransferase, *hph*), and the *Aspergillus nidulans trpC* terminator, using PfuUltra II Fusion HS DNA Polymerase (Stratagene). The PCR-amplified fragment was cloned into pCR-blunt II-TOPO (Invitrogen) and transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen). A colony was isolated and sequencing confirmed that the extracted plasmid displayed a mutated 5' *Asc*I restriction site, which was replaced with a *Nar*I site (GGCGCGCC→GGCGCCC). The construct was then digested with *Nar*I (Roche), *Asc*I (NEB) and *Dra*I (Roche), and the resulting 2.5 kb fragment was isolated using the QiaQuick Gel Extraction kit in accordance with the manufacturer's protocol (Qiagen) in preparation for cloning subsequently into the Bxl1-deletion plasmid.

Ligation of the two isolated fragments as described above was performed with 1 µL 10X Ligation Buffer (Roche), 1 µL 5 U/ml T4 DNA Ligase (Roche), and 50 ng of each fragment in a reaction volume of 10 µL. The ligation mixture was cloned into *E. coli One Shot® TOP10* Chemically Competent cells and a single colony was isolated. The *bxl1*-deletion vector (Figure 17, pCR-BluntII-TOPO, *bxl1* deletion, *hph-loxP*) containing the *loxP*-flanked hygromycin resistance gene was extracted from the *E.coli* and the appropriate ligations were verified by restriction digest using *Bmt*1, resulting in 4 fragments of 4754, 2195, 1899, and 1182 base pairs.

The *bxl1*-deletion cassette was generated by amplifying the fragment from plasmid pCR-BluntII-TOPO, *bxl1* deletion, *hph-loxP* using primers MH379/ MH380 in a total volume of 10 mL and PfuUltra II Fusion HS DNA Polymerase (Stratagene). The PCR product was cleaned and concentrated using a QiaexII kit (Qiagen). The DNA was further concentrated by SpeedVac to a concentration of about 1.5 mg/mL.

10.2 Transformation of bxl1 deletion plasmid into *T. reesei*

The DNA from the *bxl1*-deletion cassette was transformed into *T. reesei* strain #229, which overexpressed *T. reesei* Bgl1 and *T. reesei* Xyn3 as previously described herein. Transformants were selected on a medium containing 100 ppm Hygromycin B (Invitrogen). A transformant containing the *bxl1* deletion was selected by PCR. A *bxl1*-deficient *T. reesei* strain is referred to herein as a "*bxl1*-" strain.

11. EXAMPLE 5: EXPRESSION OF FV43D BY *T. REESEI* STRAIN

The following example shows how *Trichoderma reesei* was engineered to express Fv43D. A *T. reesei* strain engineered to express Fv43D is referred to herein as a "Fv43D+" strain

11.1 Construction of expression cassette

A *F. verticillioides β-xylosidase* Fv43D expression cassette was constructed by PCR amplification of the Fv43D gene from *F.verticillioides* genomic DNA sample using the primers SK1322/SK1297 (Table 5). A region of the promoter of the endoglucanase gene *egl1* was amplified by PCR from a *T. reesei* genomic DNA sample extracted from an engineered *T. reesei* strain RL-P37 (*see, e.g.*, Sheir-Neiss G. and B.S. Montenecourt, Appl. Microbiol. Biotechnology, 20 (1984) pp. 46-53), using the primers SK1236/SK1321. These two amplified fragments were subsequently fused together in a fusion PCR reaction using the primers SK1236/SK1297. The resulting fusion PCR fragment was cloned into pCR-Blunt II-TOPO vector (Invitrogen) to yield plasmid TOPO Blunt/Pegl1-Fv43D (Figure 18), which was in turn used to transform *E. coli One Shot® TOP10* chemically competent cells (Invitrogen). Plasmid DNA was extracted from several *E.coli* clones and confirmed by restriction digests.

The expression cassette from TOPO Blunt/Pegl1-Fv43D was amplified by PCR using primers SK1236/SK1297 (Table 5) to generate a DNA product for transformation of *T. reesei*. The expression cassette was co-transformed with an existing selection marker cassette containing the *als* gene (acetolactate synthase).

11.2 Transformation of *bxl1*-deficient *T. reesei* with Fv43D expression cassette

The *bxl1*-deletion *T. reesei* host strain is transformed with the *Fv43D* expression cassette (comprising an *eg1* promoter, an Fv43D open reading frame, and a native terminator of Fv43D) and an existing selection marker cassette containing the native *als* gene, using a standard transformation method such as, for example, electroporation (*see, e.g.*, PCT patent application publication WO 08/153712). Transformants are selected on minimal media agar plates containing chlorimuron ethyl. These transformants are bxl1$^-$Fv43D$^+$ *T. reesei*.

11.3 Transformation of *bxl1*-expressing *T. reesei* with Fv43D expression cassette

A *T. reesei* host strain having a wild type *bxl1* gene (*see*, PCT patent application publication WO 2005/001036 A2) wastransformed with the *Fv43D* expression cassette

(comprising an *eg1* promoter, an Fv43D open reading frame, and a native Fv43D terminator) and an existing selection marker cassette containing the native *als* gene, using standard transformation methods such as, for example, electroporation (*see, e.g.,* PCT patent application publication WO 08/153712). Transformants were selected on minimal media agar plates containing chlorimuron ethyl. These transformants are Fv43D^{+} *T. reesei.*

12. EXAMPLE 6: USE OF CELLULASE PRODUCED ENGINEERED *T. REESEI* STRAINS IN SSF

The transformants *bxl1*$^{-}$Fv43D^{+} *T. reesei,* and Fv43D^{+} *T. reesei* are used to produce cellulase-containing culture broths. These culture broths are then used in SSf reactions in continuous, batch, or fed-batch configurations, as described herein to reduce the production of AXP and reduce the sugar yield loss by the production of a trans-xylosication of reverse hydrolysis product.

In a particular example, the transformant of a *bxl1*$^{-}$ *T. reesei* strain #229 ("229 Bxl del") was cultured to produce a cellulase-containing culture broth, which was then supplemented with either a purified *T. reesei* Bxl1 (at 0.5 mg/g or 1 mg/g of xylan) or a purified FV43D, at 1 mg/g of xylan. The concentration of the EXP and ethanol at days 1 and 3 were plotted in Figures 43A (Day 1) and 43B (Day 3), respectively. The reactions were carried out under the recombinant *Zymomonas* SSF conditions using a 25 wt.% solids loading of dilute ammonia pretreated corn cob substrate.

Any prior art reference or statement provided in the specification is not to be taken as an admission that such art constitutes, or is to be understood as constituting, part of the common general knowledge.

The Claims defining the invention are as follows:

1. A method for simultaneous saccharification and fermentation (SSF) comprising culturing a complete fermentation medium, said complete fermentation medium comprising at least one fermenting microorganism, at least one xylan-containing biomass, at least one cellulase, at least one hemicellulase, at least one retaining β-xylosidase, and at least one inverting β-xylosidase, for a period and under conditions suitable for producing a fermentation product, wherein the inverting β-xylosidase is an Fv43D, a Pf43A, an Fv43E, an Fv43B, an Af43A, an Fo43A, a Gz43A, or an XynB3 polypeptide, and the complete fermentation medium comprises a greater amount of inverting β-xylosidases than that of retaining β-xylosidases on a mole basis, molecular weight basis, or on both a mole basis and a molecular weight basis.

2. The method of claim 1, wherein the complete fermentation medium comprises an effective amount of the inverting β-xylosidase such that the complete fermentation medium produces less short chain alkyl-β-xylopyranoside ("AXP") than does a control fermentation medium lacking the inverting β-xylosidase.

3. The method of claim 1 or 2, wherein the complete fermentation medium comprises an effective amount of the inverting β-xylosidase such that the complete fermentation medium produces at least 40% less AXP than does a control fermentation medium lacking the inverting β-xylosidase.

4. The method of any one of claims 1 to 3, wherein the AXP is a methyl-β-xylopyranoside (MXP), an ethyl-β-xylopyranoside (EXP), a propyl-β-xylopyranoside (PXP), or a butyl-β-xylopyranoside (BXP).

5. The method of any one of claims 1 to 4, wherein the complete fermentation medium comprises an effective amount of the inverting β-xylosidase to increase the yield of the fermentation product, as compared to the yield of the fermentation product from culturing a control fermentation medium lacking the inverting β-xylosidase.

6. The method of claim 5, wherein the yield of the fermentation product is increased by at least 1%.

7. The method of claim 5, wherein the yield of fermentation product is increased by at least 5%.

8. The method of claim 5, wherein the yield of fermentation product is increased by at least 10%.

9. The method of any one of claims 1 to 8, wherein the fermentation product is an alcohol.

10. The method of claim 9, wherein the alcohol is methanol, ethanol, propanol, propane-1,3-diol, or butanol.

11. The method of claim 1, wherein the Fv43D polypeptide has at least 90% sequence identity to an amino acid sequence corresponding to SEQ ID NO:2, or to residues 21 to 350 of SEQ ID NO:2; the Pf43A polypeptide has at least 90% sequence identity an amino acid sequence corresponding to SEQ ID NO:8, or to residues 21 to 445 of SEQ ID NO:8; the Fv43E polypeptide has at least 90% sequence identity to an amino acid sequence corresponding to SEQ ID NO:10, or to residues 19 to 530 of SEQ ID NO:10; the Fv43B polypeptide has at least 90% sequence identity to an amino acid sequence corresponding to SEQ ID NO:12, or to residues 17 to 574 of SEQ ID NO:12; the Af43A polypeptide has at least 90% sequence identity to an amino acid sequence corresponding to SEQ ID NO:14, or to residues 15-558 of SEQ ID NO:14; the Fo43A polypeptide has at least 90% sequence identity to an amino acid sequence corresponding to SEQ ID NO:24, or to residues 21-348 of SEQ ID NO:24; the Gz43A polypeptide has at least 90% sequence identity to an amino acid sequence corresponding to SEQ ID NO:22, or to residues 19-340 of SEQ ID NO:22; or the XynB3 polypeptide has at least 90% sequence identity to an amino acid sequence corresponding to SEQ ID NO:25.

12. The method of any one of the preceding claims, wherein the inverting β-xylosidase is present in the complete fermentation medium at a concentration of 0.3 mg to 10 mg per gram of xylan in the xylan-containing biomass.

13. The method of any one of the preceding claims, which is performed as a continuous, a batch, or a fed-batch SSF process.

14. The method of any one of the preceding claims, wherein the fermenting microorganism is *Saccharomyces cerevisiae* or *Zymomonas mobilis*.

15. The method of any one of the preceding claims, wherein the xylan-containing biomass is corn stover, bagasses, sorghum, giant reed, elephant grass, miscanthus, Japanese cedar, wheat straw, switchgrass, hardwood pulp, or softwood pulp.

16. The method of any one of the preceding claims, wherein the xylan-containing biomass has been pretreated.

17. The method of any one of the preceding claims, wherein the ratio of inverting β-xylosidases to retaining β-xylosidases in the complete fermentation medium is at least 2:1, on a mole basis, molecular weight basis, or on both a mole basis and molecular weight basis.

Process	Substrate	Solids loading	Enzymes used	EtOH at day 3 of fermentation (g/L)	EXP at day 3 of fermentation (g/L)
Batch	Glucose+xylose (80 g/L+70 g/L)	~	No enzyme added	77.6	1.53
Fed-batch	Glucose+xylose (80 g/L+70 g/L)	~	No enzyme added	78.0	0.99
SHF	Corn cob	25%	A1500+MF (20+5 mg/g)	32.1	1.11
SSF	Corn cob	25%	A1500+MF (20+5 mg/g)	37.5	6.10
Fed-batch SSF	Corn cob	25%	A1500+MF (20+5 mg/g)	48.0	14.3

TABLE 1

Process	Substrate	Substrate's xylan composition	Solids loading	Enzymes used	EtOH at day 3 of fermentation (g/L)	EXP at day 3 of fermentation (g/L)
SSF	Bagasse	3.1%	20%	A1500+MF (20+5 mg/g)	20.9	0.91
SSF	Softwood pulp	8.0%	10%	A1500+MF (20+5 mg/g)	48.6	4.75
SHF	Mixed hardwood pulp	20.7%	10%	A1500+MF (20+5 mg/g)	41.6	3.52
SSF	Mixed hardwood pulp	20.7%	10%	A1500+MF (20+5 mg/g)	43.7	10.7
SSF	Mixed hardwood pulp	20.7%	12%	A1500+MF (20+5 mg/g)	60.2	16.1

TABLE 2

Time (h)	MF EXP	MF Xylose	Fv43D Exp	Fv43D Xylose	Fv3A EXP	Fv3A Xylose
0	2778	12043	0	7046	3991	11900
0.5	47400	91600	3056	59800	44000	114000
1	72000	125000	4563	84200	62500	147500
2	98300	151000	9534	117800	81420	178000
3.5	98800	153000	8359	118000	82100	180000
20	140000	220000	32800	326000	94100	264000
	Refractive Index Area					

TABLE 3

Primer sequences for the construction of *bxl1* deletion cassette			
Primer Name	**Seq ID No.**	**Primer sequence (5'→3')**	**Description**
MH290	27	caaGGCGCGCCaagtATAACTTCGTATAATGTATGCTATACGAAGTTATCGGCCGGCGTATTGGGTGTTACG	hph reverse primer AscI loxP site
MH292	28	GAAGGCGCGCCACAGATAACTTCGTATAGCATACATTATACGAAGTTATcctgggcttgtgactggtcgcgag	hph forward primer full promoter AscI loxP site
MH375	29	ccatgtcacctgtcttgaacac	Bxl1 5' forward
MH376	30	**caaggcgc**GCCATCTCTTTCGATCTCAACAG	Bxl1 5' reverse AscI
MH377	31	gattgcgatcgccgtctacaacgttttcaacc	Bxl1 3' forward AsiSI AclI
MH378	32	GGTCCAACCTTGAATGTAACAGC	Bxl1 3' reverse primer
MH379	33	gtgtcgctgaacataaggtctc	Bxl1 deletion nested forward primer
MH380	34	CCTCCATTCTTCCAACAAGCC	Bxl1 deletion nested reverse primer

TABLE 4

Primer sequences for the construction of *F. verticillioides β-xylosidase* Fv43D expression cassette		
Primer Name	**Primer sequence (5'→3')**	**SEQ ID NO.**
Forward Primer (SK1322)	CACCATGCAGCTCAAGTTTCTGTC	35
Reverse Primer (SK1297)	GGTTACTAGTCAACTGCCCGTTCTGTAGCGAG	36
Forward Primer (SK1236)	CATGCGATCGCGACGTTTTGGTCAGGTCG	37
Reverse Primer (SK1321)	GACAGAAACTTGAGCTGCATGGTGTGGGACAACAAGAAGG	38

TABLE 5

SEQ ID NO:	Nucleotide or amino acid	Description
1.	Nucleotide	Nucleotide sequence for Fv43D, a GH43 family enzyme from *Fusarium verticillioides*
2.	Amino acid	Protein sequence of Fv43D
3.	Nucleotide	Nucleotide sequence of Bxl1, a GH3 β-xylosidase from *Trichoderma reesei*
4.	Amino acid	Protein sequence of Bxl1
5.	Nucleotide	Nucleotide sequence for Fv3A, a GH3 family enzyme from *Fusarium verticillioides*
6.	Amino acid	Protein sequence of Fv3A
7.	Nucleotide	Nucleotide sequence for Pf43A, a GH43 family enzyme from *Penicillium funiculosum*
8.	Amino acid	Protein sequence of Pf43A
9.	Nucleotide	Nucleotide sequence for Fv43E, a GH43 family enzyme from *Fusarium verticillioides*
10.	Amino acid	Protein sequence of Fv43E
11.	Nucleotide	Nucleotide sequence for Fv43B, a GH43 family enzyme from *Fusarium verticillioides*
12.	Amino acid	Protein sequence of Fv43B
13.	Nucleotide	Nucleotide sequence for Af43A, a GH43 family enzyme from *Aspergillus fumigatus*
14.	Amino acid	Protein sequence of Af43A
15.	Nucleotide	Nucleotide sequence for Fv51A, a GH51 family enzyme from *Fusarium vertiilloides*
16.	Amino acid	Protein sequence of Fv51A
17.	Nucleotide	Nucleotide sequence for Xyn3, a GH10 family xylanase from *Trichoderma reesei*
18.	Amino acid	Protein sequence of Xyn3
19.	Nucleotide	Nucleotide sequence for xlnA, an *Aspergillus tubingensis* xylanase
20.	Amino acid	xlnA protein sequence
21.	Nucleotide	Nucleotide sequence for Gz3A, a GH43 family enzyme from *Gibberella zeae*
22.	Amino acid	Protein sequence of Gz3A
23.	Nucleotide	Nucleotide sequence for Fo43A, a GH43 family enzyme from *Fusarium oxysporum*
24.	Amino acid	Protein sequence of Fo43A
25.	Amino acid	Protein sequence of XynB3, a GH43 family enzyme from *Geobacillus stearothermophilus* T-6
26.	Amino acid	Protein sequence of Bgl1, a GH3 family enzyme from *T. reesei*
39.	Nucleotide	Nucleotide sequence for XlnD, a GH3 family enzyme from *A. japonicus*
40.	Amino acid	Protein sequence of XlnD
41.	Nucleotide	Nucleotide sequence of Fv30A, a GH30 family enzyme from *F. verticillioides*
42.	Amino acid	Protein sequence of Fv30A
43.	Nucleotide	Nucleotide sequence of Fv30B, a GH30 family enzyme

		from *F. verticillioides*
44.	Amino acid	Protein sequence of Fv30B
45.	Nucleotide	Nucleotide sequence of Fv39A, a GH39 family enzyme from *F. verticillioides*
46.	Amino acid	Protein sequence of Fv39A
47.	Nucleotide	Nucleotide sequence of Fv39B, a GH39 family enzyme from *F. verticillioides*
48.	Amino acid	Protein sequence of Fv39B
49.	Nucleotide	Nucleotide sequence of XynB, a GH39 family enzyme from *T. saccarolyticum*
50.	Amino acid	Protein sequence of XynB
51.	Nucleotide	Nucleotide sequence of XylA, a GH52 family enzyme from *G. stearothermophilus*
52.	Amino acid	Protein sequence of XylA
53.	Nucleotide	Nucleotide sequence of Xyl1, a GH54 family enzyme from *T. koningii* (*H. Koningii*)
54.	Amino acid	Protein sequence of Xyl1

TABLE 6

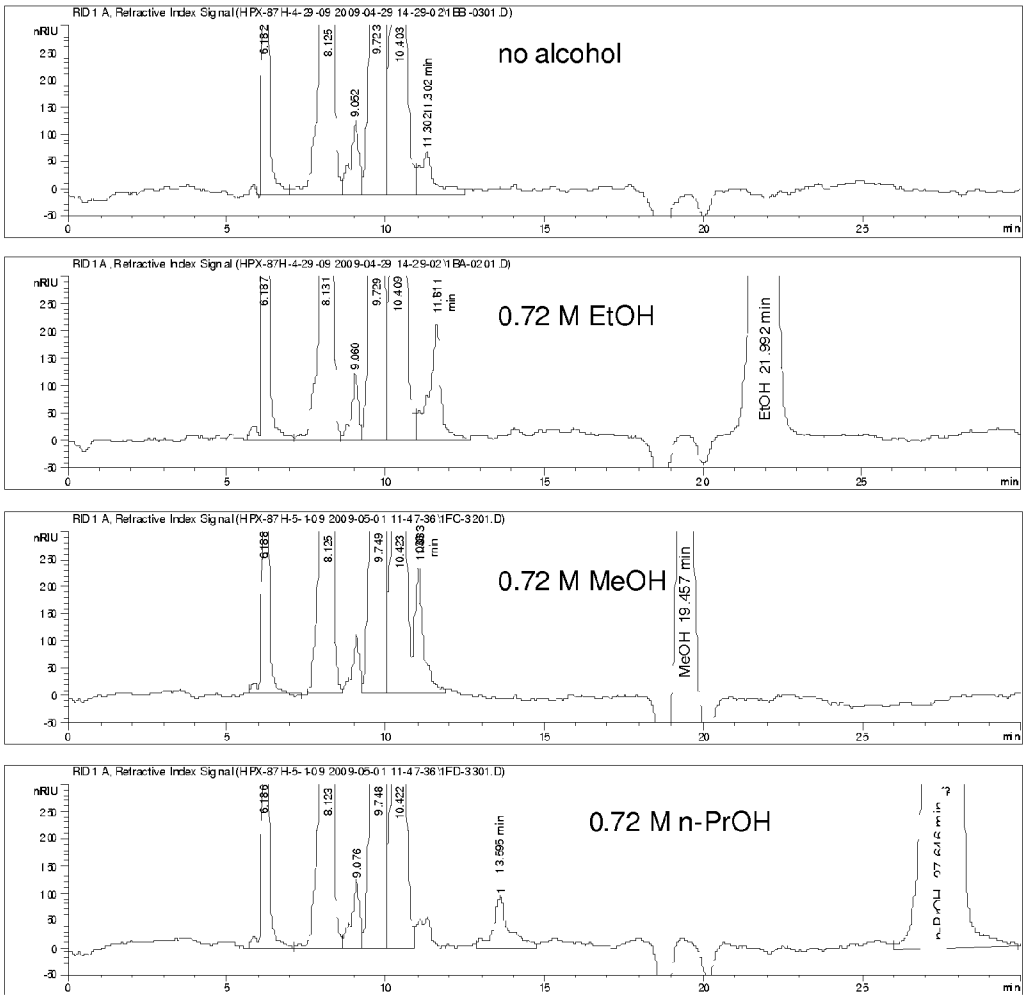

no alcohol
0.72 M EtOH
EtOH 21.992 min
0.72 M MeOH
MeOH 19.457 min
0.72 M n-PrOH
13.595 min
nRIU
min

FIGURE 1

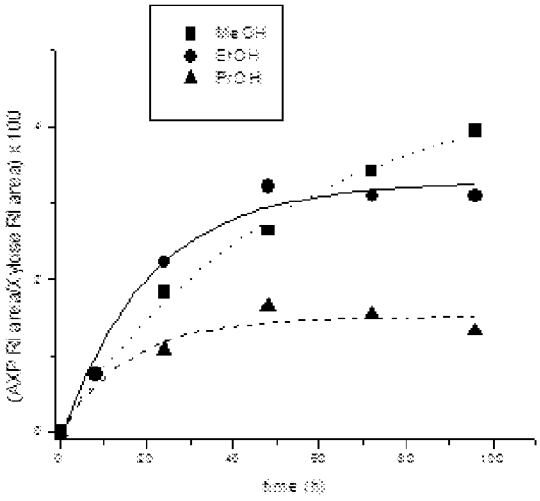


FIGURE 2

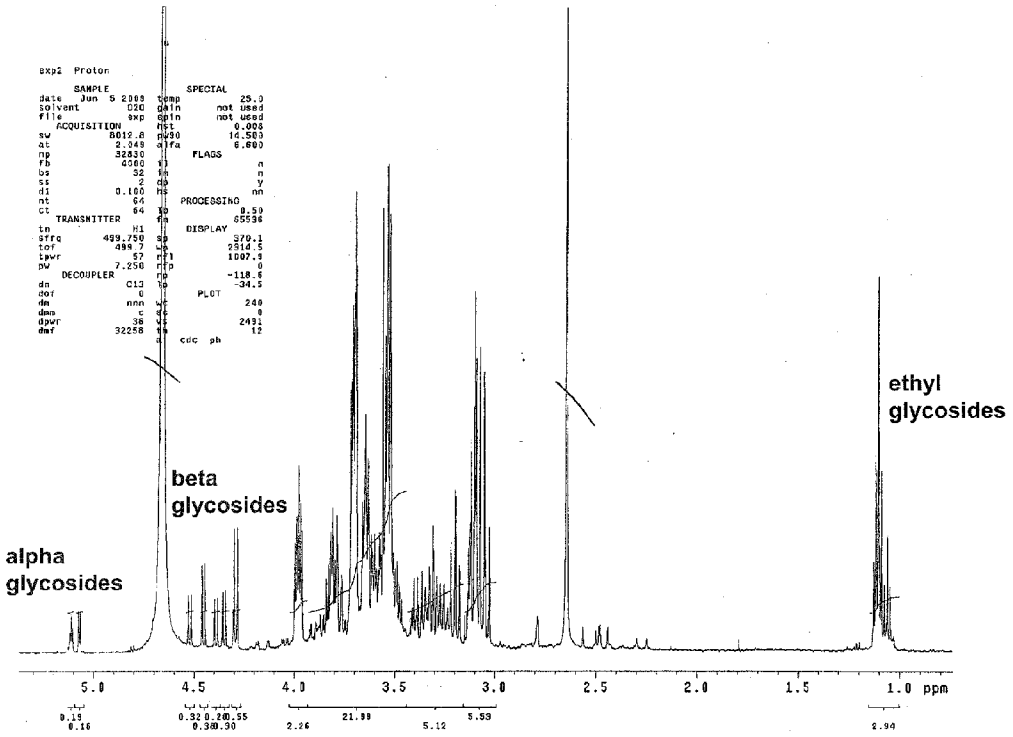

alpha
glycosides
beta
glycosides
ethyl
glycosides
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0 ppm

FIGURE 3

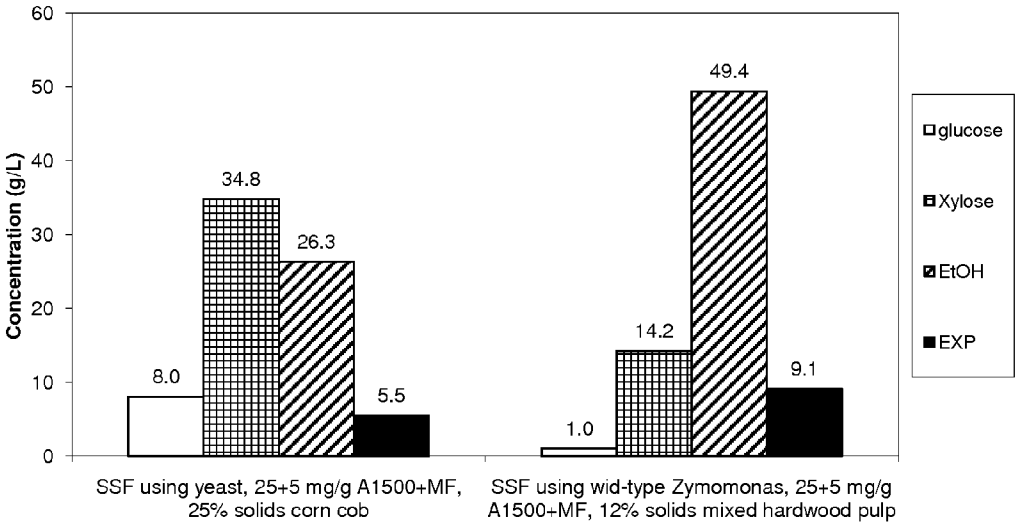

60
50
40
30
20
10
0
Concentration (g/L)
8.0
34.8
26.3
5.5
1.0
14.2
49.4
9.1
glucose
Xylose
EtOH
EXP
SSF using yeast, 25+5 mg/g A1500+MF, 25% solids corn cob
SSF using wid-type Zymomonas, 25+5 mg/g A1500+MF, 12% solids mixed hardwood pulp

FIGURE 4

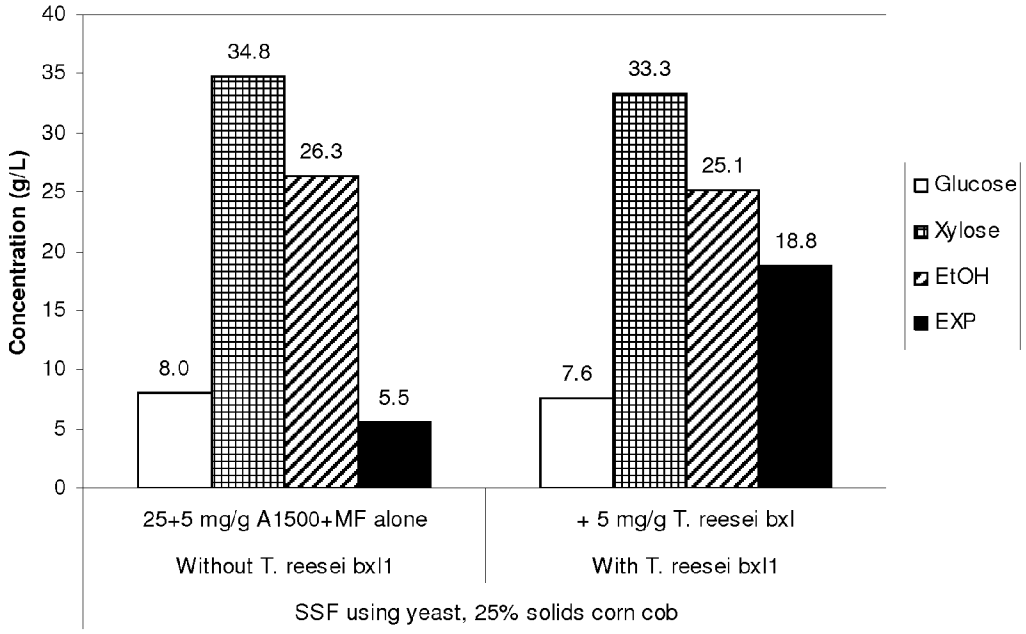

40
35
30
25
20
15
10
5
0
Concentration (g/L)
34.8
26.3
8.0
5.5
33.3
25.1
18.8
7.6
Glucose
Xylose
EtOH
EXP
25+5 mg/g A1500+MF alone
Without T. reesei bxl1
+ 5 mg/g T. reesei bxl
With T. reesei bxl1
SSF using yeast, 25% solids corn cob

FIGURE 5

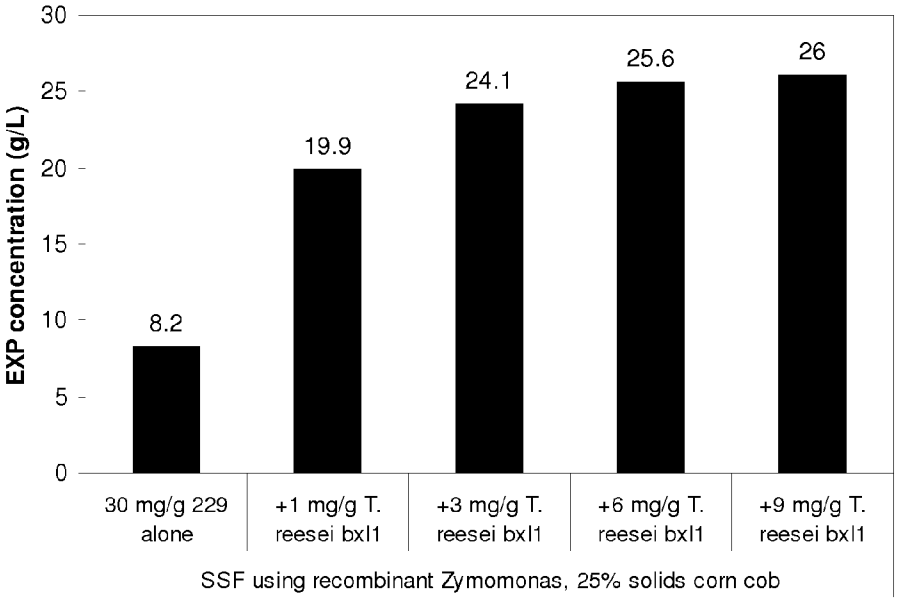

30
25
20
15
10
5
0
EXP concentration (g/L)
8.2
19.9
24.1
25.6
26
30 mg/g 229 alone
+1 mg/g T. reesei bxl1
+3 mg/g T. reesei bxl1
+6 mg/g T. reesei bxl1
+9 mg/g T. reesei bxl1
SSF using recombinant Zymomonas, 25% solids corn cob

FIGURE 6

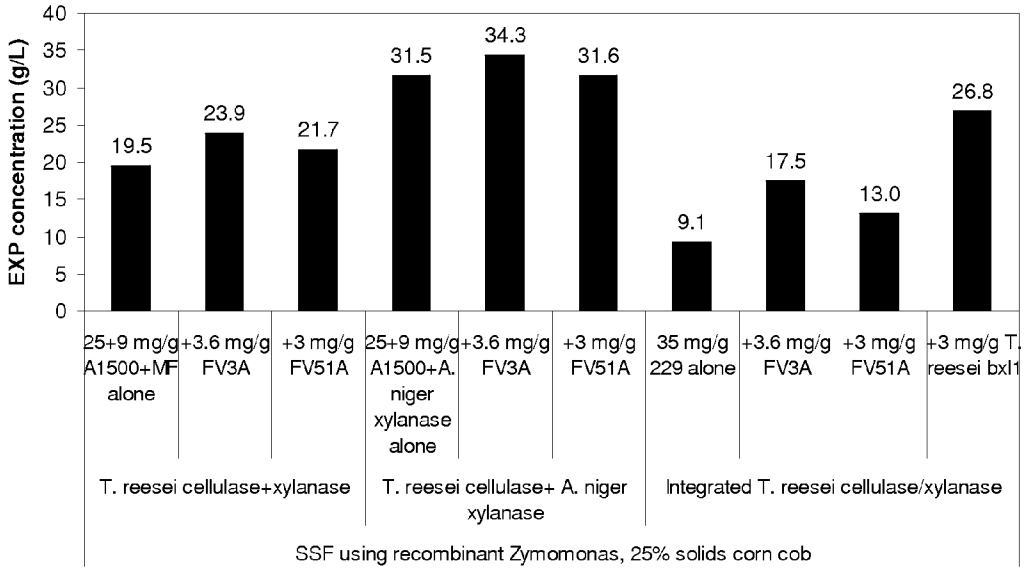

EXP concentration (g/L)
40
35
30
25
20
15
10
5
0
19.5
23.9
21.7
31.5
34.3
31.6
9.1
17.5
13.0
26.8
25+9 mg/g A1500+MF alone
+3.6 mg/g FV3A
+3 mg/g FV51A
25+9 mg/g A1500+A. niger xylanase alone
+3.6 mg/g FV3A
+3 mg/g FV51A
35 mg/g 229 alone
+3.6 mg/g FV3A
+3 mg/g FV51A
+3 mg/g T. reesei bxl1
T. reesei cellulase+xylanase
T. reesei cellulase+ A. niger xylanase
Integrated T. reesei cellulase/xylanase
SSF using recombinant Zymomonas, 25% solids corn cob

FIGURE 7

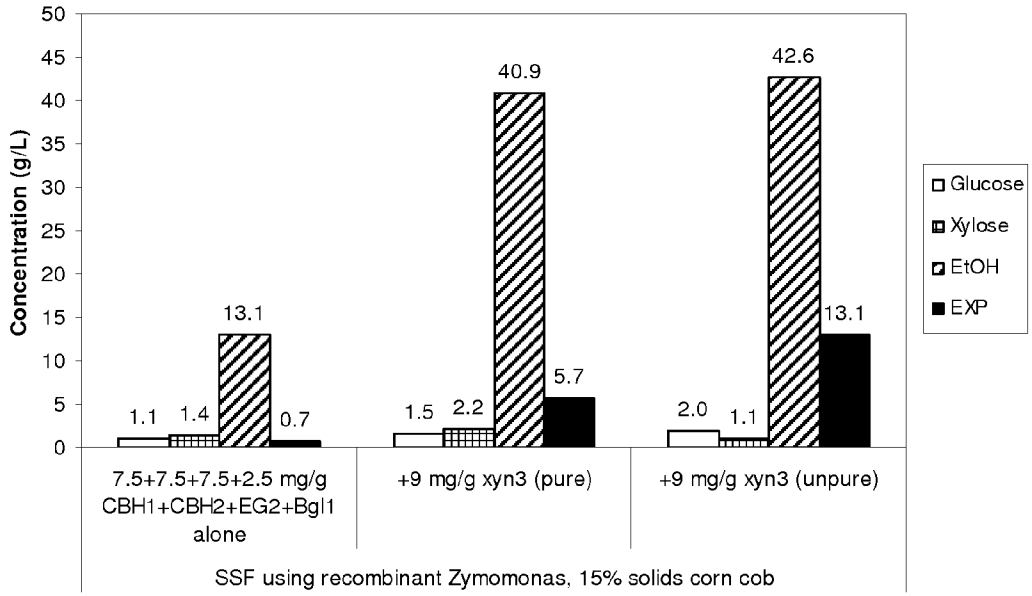

Concentration (g/L)
50
45
40
35
30
25
20
15
10
5
0
1.1
1.4
13.1
0.7
1.5
2.2
40.9
5.7
2.0
1.1
42.6
13.1
Glucose
Xylose
EtOH
EXP
7.5+7.5+7.5+2.5 mg/g CBH1+CBH2+EG2+Bgl1 alone
+9 mg/g xyn3 (pure)
+9 mg/g xyn3 (unpure)
SSF using recombinant Zymomonas, 15% solids corn cob

FIGURE 8

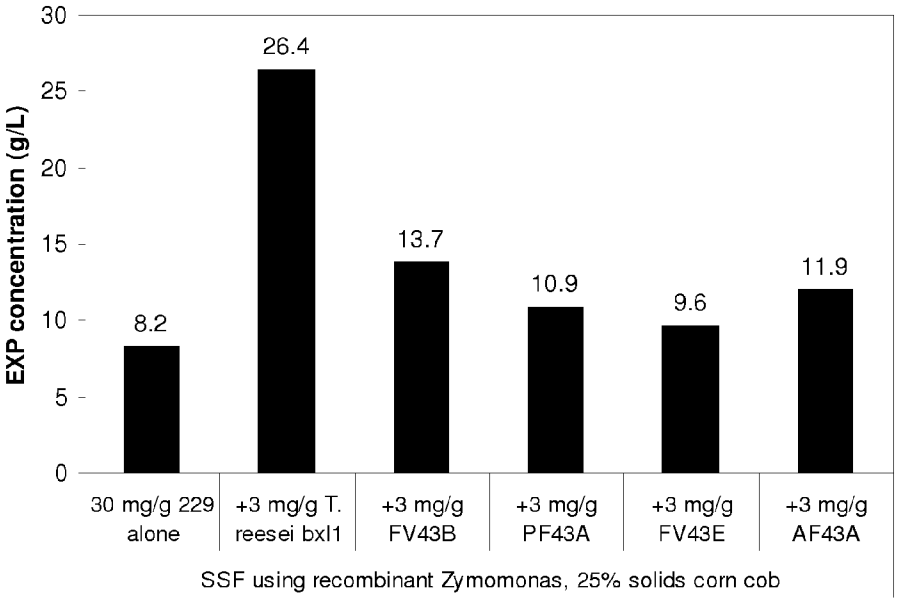

30
25
20
15
10
5
0
EXP concentration (g/L)
8.2
26.4
13.7
10.9
9.6
11.9
30 mg/g 229 alone
+3 mg/g T. reesei bxl1
+3 mg/g FV43B
+3 mg/g PF43A
+3 mg/g FV43E
+3 mg/g AF43A
SSF using recombinant Zymomonas, 25% solids corn cob

FIGURE 9

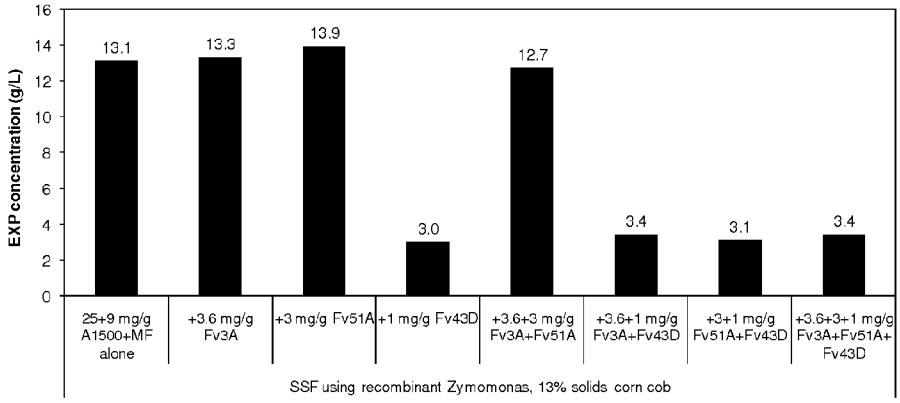

EXP concentration (g/L)
16
14
12
10
8
6
4
2
0
13.1
13.3
13.9
3.0
12.7
3.4
3.1
3.4
25+9 mg/g A1500+MF alone
+3.6 mg/g Fv3A
+3 mg/g Fv51A
+1 mg/g Fv43D
+3.6+3 mg/g Fv3A+Fv51A
+3.6+1 mg/g Fv3A+Fv43D
+3+1 mg/g Fv51A+Fv43D
+3.6+3+1 mg/g Fv3A+Fv51A+ Fv43D
SSF using recombinant Zymomonas, 13% solids corn cob

FIGURE 10

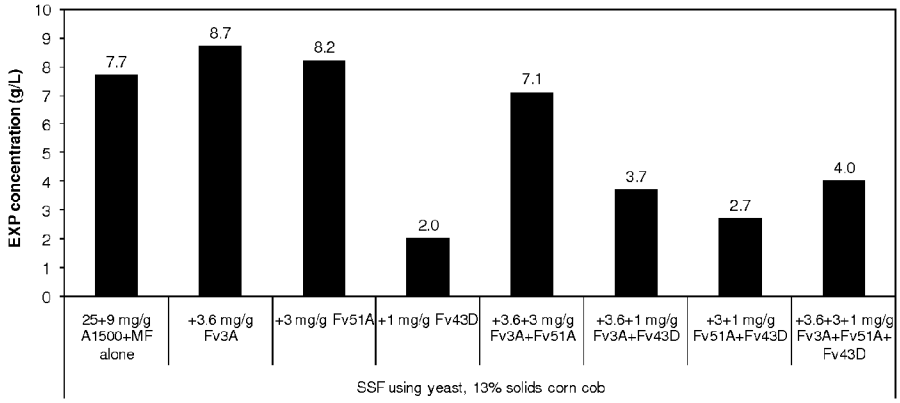

EXP concentration (g/L)
10
9
8
7
6
5
4
3
2
1
0
7.7
8.7
8.2
2.0
7.1
3.7
2.7
4.0
25+9 mg/g A1500+MF alone
+3.6 mg/g Fv3A
+3 mg/g Fv51A
+1 mg/g Fv43D
+3.6+3 mg/g Fv3A+Fv51A
+3.6+1 mg/g Fv3A+Fv43D
+3+1 mg/g Fv51A+Fv43D
+3.6+3+1 mg/g Fv3A+Fv51A+ Fv43D
SSF using yeast, 13% solids corn cob

FIGURE 11

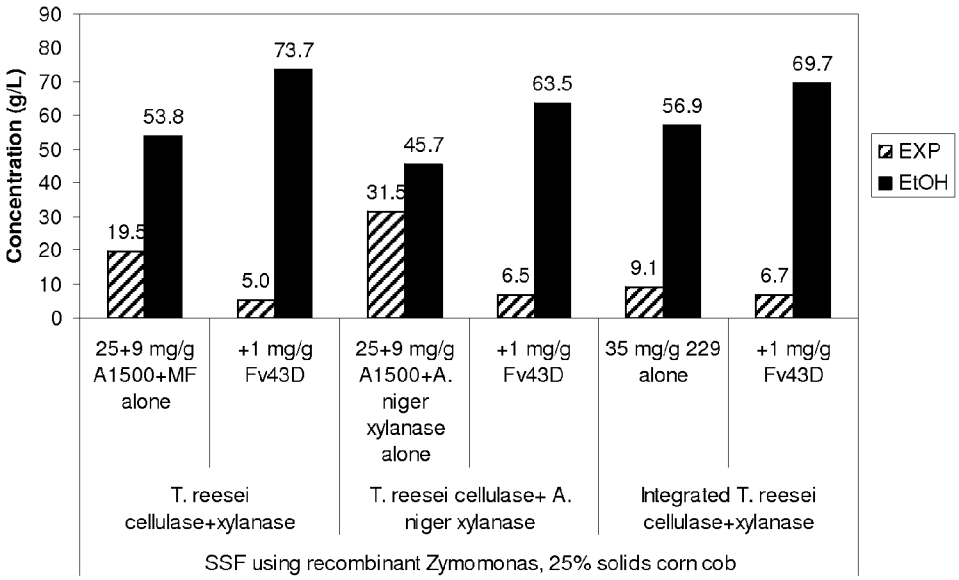

Concentration (g/L)
90
80
70
60
50
40
30
20
10
0
19.5
53.8
5.0
73.7
31.5
45.7
6.5
63.5
9.1
56.9
6.7
69.7
EXP
EtOH
25+9 mg/g A1500+MF alone
+1 mg/g Fv43D
25+9 mg/g A1500+A. niger xylanase alone
+1 mg/g Fv43D
35 mg/g 229 alone
+1 mg/g Fv43D
T. reesei cellulase+xylanase
T. reesei cellulase+ A. niger xylanase
Integrated T. reesei cellulase+xylanase
SSF using recombinant Zymomonas, 25% solids corn cob

FIGURE 12

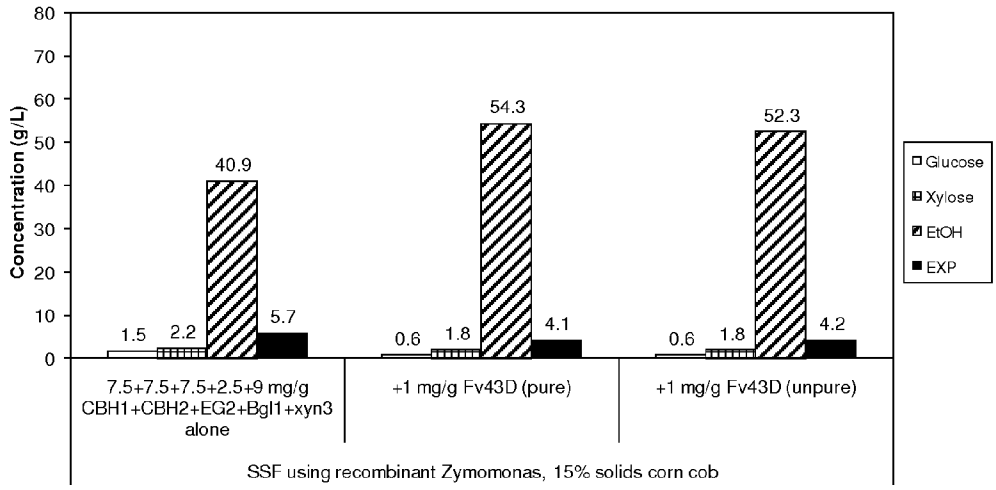

80
70
60
50
40
30
20
10
0
Concentration (g/L)
1.5
2.2
40.9
5.7
0.6
1.8
54.3
4.1
0.6
1.8
52.3
4.2
Glucose
Xylose
EtOH
EXP
7.5+7.5+7.5+2.5+9 mg/g CBH1+CBH2+EG2+Bgl1+xyn3 alone
+1 mg/g Fv43D (pure)
+1 mg/g Fv43D (unpure)
SSF using recombinant Zymomonas, 15% solids corn cob

FIGURE 13

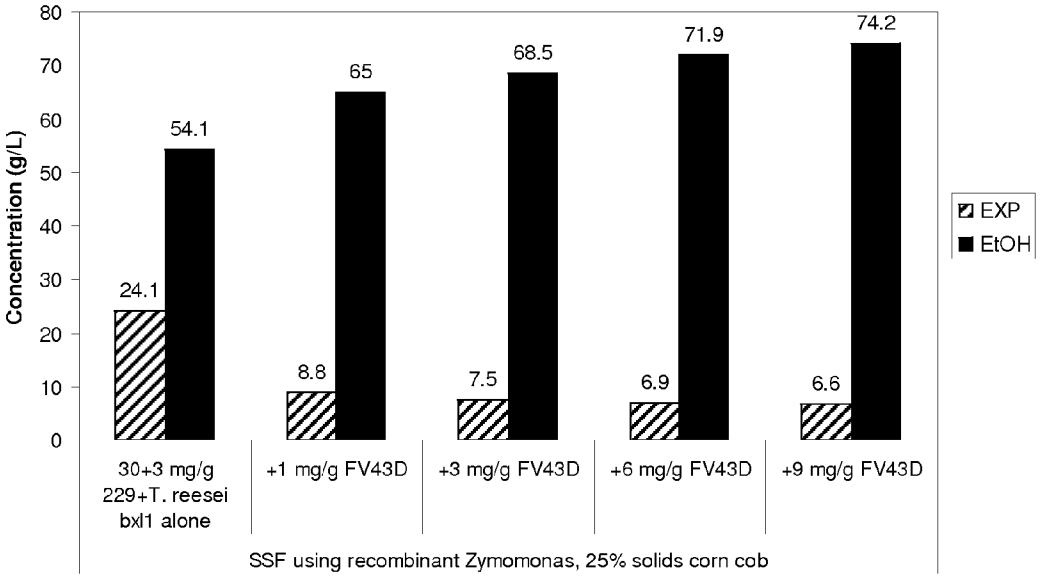


FIGURE 14A

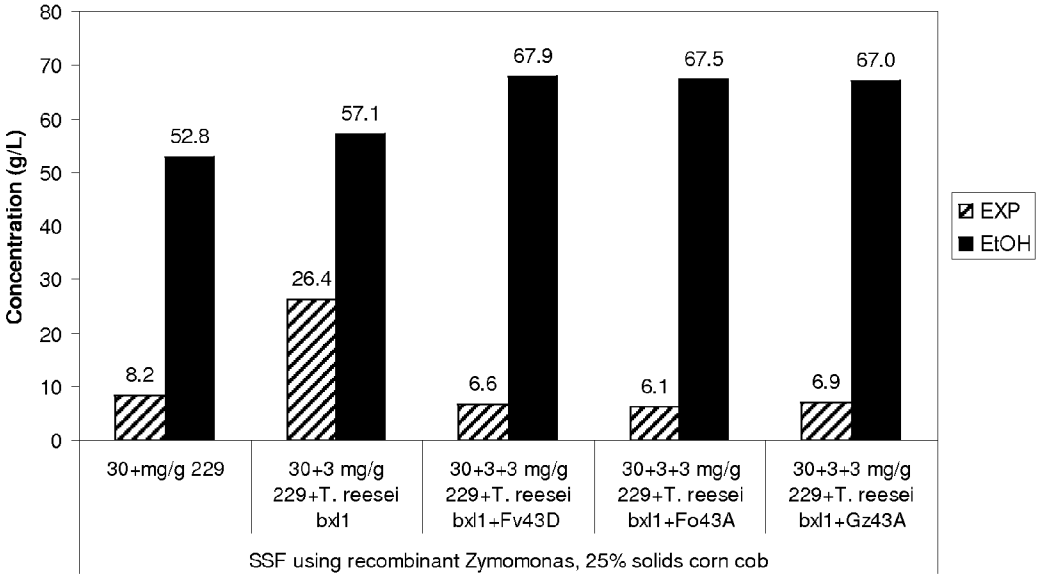


FIGURE 14B

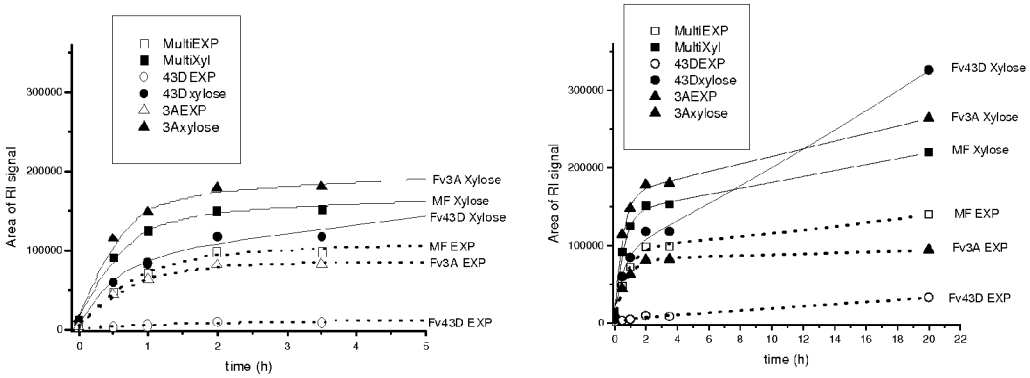

MultiEXP
MultiXyl
43DEXP
43Dxylose
3AEXP
3Axylose
Area of RI signal
time (h)
Fv3A Xylose
MF Xylose
Fv43D Xylose
MF EXP
Fv3A EXP
Fv43D EXP

FIGURE 15

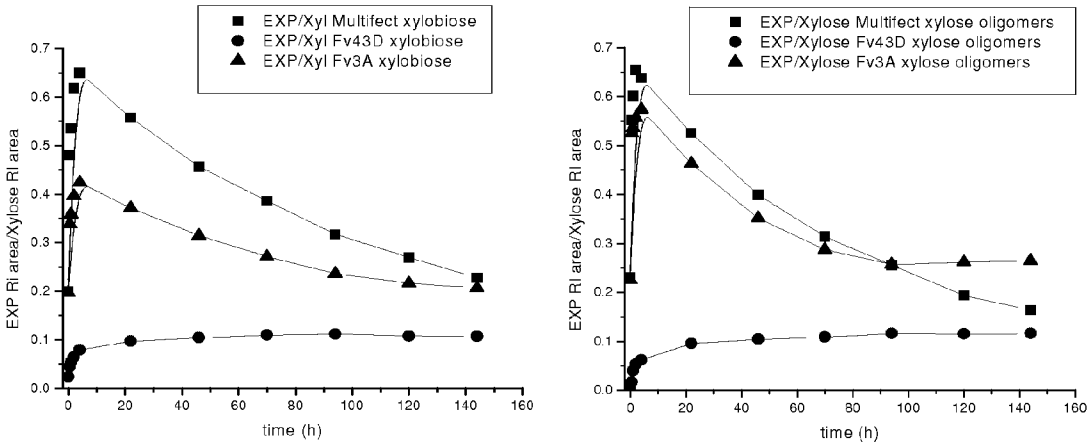

EXP/Xyl Multifect xylobiose
EXP/Xyl Fv43D xylobiose
EXP/Xyl Fv3A xylobiose
EXP RI area/Xylose RI area
time (h)
EXP/Xylose Multifect xylose oligomers
EXP/Xylose Fv43D xylose oligomers
EXP/Xylose Fv3A xylose oligomers
EXP RI area/Xylose RI area
time (h)

FIGURE 16

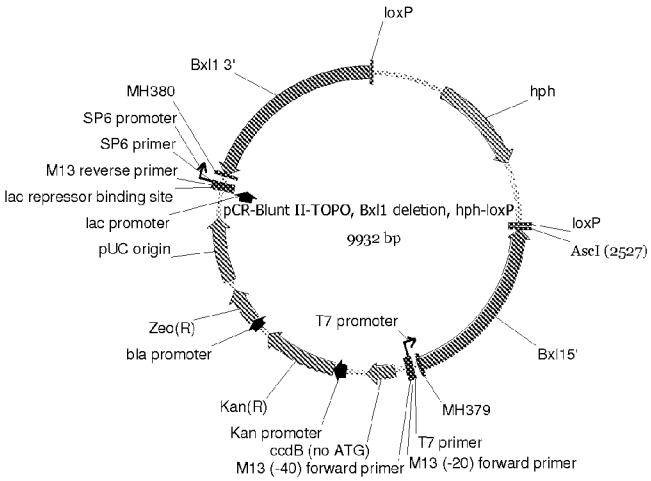

loxP
Bxl1 3'
MH380
SP6 promoter
SP6 primer
M13 reverse primer
lac repressor binding site
lac promoter
pUC origin
pCR-Blunt II-TOPO, Bxl1 deletion, hph-loxP
9932 bp
hph
loxP
AscI (2527)
Bxl15'
T7 promoter
Zeo(R)
bla promoter
Kan(R)
Kan promoter
ccdB (no ATG)
M13 (-40) forward primer
MH379
T7 primer
M13 (-20) forward primer

FIGURE 17

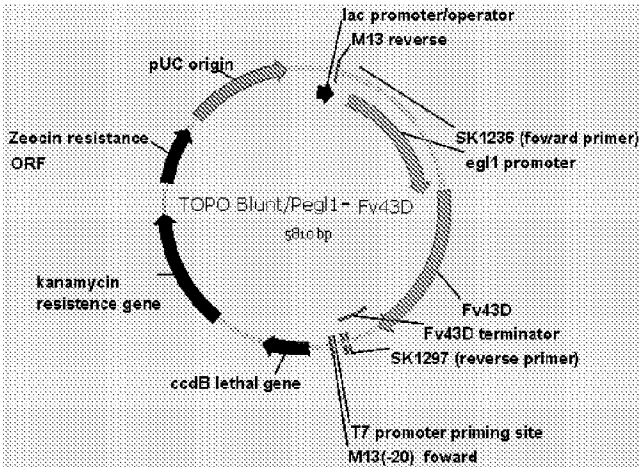

lac promoter/operator
M13 reverse
pUC origin
Zeocin resistance
ORF
SK1236 (foward primer)
egl1 promoter
TOPO Blunt/Pegl1 - Fv43D
5810 bp
kanamycin
resistence gene
Fv43D
Fv43D terminator
SK1297 (reverse primer)
ccdB lethal gene
T7 promoter priming site
M13(-20) foward

FIGURE 18

SEQ ID NO:1
Nucleotide sequence for Fv43D, GH43D family enzyme from *Fusarium verticillioides*

atgcagctcaagtttctgtcttcagcattgttgctgtctttgaccggcaattgcgctgcgcaagacac
taatgatatccctcctctgatcaccgacctctggtctgcggatccctcggctcatgttttcgagggca
aactctgggtttacccatctcacgacatcgaagccaatgtcgtcaacggcaccggaggcgctcagtac
gccatgagagattatcacacctattccatgaagaccatctatggaaaagatcccgttatcgaccatgg
cgtcgctctgtcagtcgatgatgtcccatgggccaagcagcaaatgtgggctcctgacgcagcttaca
agaacggcaaatattatctctacttccccgccaaggataaagatgagatcttcagaattggagttgct
gtctccaacaagcccagcggtcctttcaaggccgacaagagctggatccccggtacttacagtatcga
tcctgctagctatgtcgacactaatggcgaggcatacctcatctggggcggtatctggggcggccagc
ttcaggcctggcaggatcacaagacctttaatgagtcgtggctcggcgacaaagctgctcccaacggc
accaacgccctatctcctcagatcgccaagctaagcaaggacatgcacaagatcaccgagacaccccg
cgatctcgtcatcctggcccccgagacaggcaagccccttcaagcagaggacaataagcgacgatttt
tcgaggggccctgggttcacaagcgcggcaagctgtactacctcatgtactctaccggcgacacgcac
ttcctcgtctacgcgacttccaagaacatctacggtccttatacctatcagggcaagattctcgaccc
tgttgatgggtggactacgcatggaagtattgttgagtacaagggacagtggtggttgttctttgcgg
atgcgcatacttctggaaaggattatctgagacaggttaaggcgaggaagatctggtatgacaaggat
ggcaagattttgcttactcgtcctaagatttag

FIGURE 19A

SEQ ID NO:2
Protein sequence of Fv43D

mqlkflssalllsltgncaaqdtndipplitdlwsadpsahvfegklwvypshdieanvvngtggaqy
amrdyhtysmktiygkdpvidhgvalsvddvpwakqqm**wapdaaykngkyylyfpakdkdeifrigva**
vsnkpsgpfkadkswipgtysidpasyvdtngeayliwggiwggqlqawqdhktfneswlgdkaapng
tnalspqiaklskdmhkitetprdlvilapetgkplqaednkrrffegpwvhkrgklyylmystgdth
flvyatskniygpytyqgkildpvdgwtthgsiveykgqwwlffadahtsgkdylrqvkarkiwydkd
gkilltrpki

FIGURE 19B

SEQ ID NO:3
Nucleotide sequence of Bxl1, a GH3 family β-xylosidase from *Trichoderma reesei*

```
atggtgaata acgcagctct tctcgccgcc ctgtcggctc tcctgcccac ggccctggcg      60
cagaacaatc aaacatacgc caactactct gctcagggcc agcctgatct ctaccccgag     120
acacttgcca cgctcacact ctcgttcccc gactgcgaac atggcccctc caagaacaat     180
ctcgtctgtg actcatcggc cggctatgta gagcgagccc aggccctcat ctcgctcttc     240
accctcgagg agctcattct caacacgcaa aactcgggcc ccggcgtgcc tcgcctgggt     300
cttccgaact accaagtctg gaatgaggct ctgcacggct tggaccgcgc caacttcgcc     360
accaagggcg gccagttcga atgggcgacc tcgttcccca tgcccatcct cactacggcg     420
gccctcaacc gcacattgat ccaccagatt gccgacatca tctcgaccca agctcgagca     480
ttcagcaaca gcggccgtta cggtctcgac gtctatgcgc caaacgtcaa tggcttccga     540
agccccctct ggggccgtgg ccaggagacg cccggcgaag acgccttttt cctcagctcc     600
gcctatactt acgagtacat cacgggcatc cagggtggcg tcgaccctga gcacctcaag     660
gttgccgcca cggtgaagca ctttgccgga tacgacctcg agaactggaa caaccagtcc     720
cgtctcggtt tcgacgccat cataactcag caggacctct ccgaatacta cactccccag     780
ttcctcgctg cggcccgtta tgcaaagtca cgcagcttga tgtgcgcata caactccgtc     840
aacggcgtgc ccagctgtgc caacagcttc ttcctgcaga cgcttttgcg cgagagctgg     900
ggcttccccg aatggggata cgtctcgtcc gattgcgatg ccgtctacaa cgttttcaac     960
cctcatgact acgccagcaa ccagtcgtca gccgccgcca gctcactgcg agccggcacc    1020
gatatcgact gcggtcagac ttacccgtgg cacctcaacg agtcctttgt ggccggcgaa    1080
gtctcccgcg gcgagatcga gcggtccgtc acccgtctgt acgccaacct cgtccgtctc    1140
ggatacttcg acaagaagaa ccagtaccgc tcgctcggtt ggaaggatgt cgtcaagact    1200
gatgcctgga acatctcgta cgaggctgct gttgagggca tcgtcctgct caagaacgat    1260
ggcactctcc ctctgtccaa gaaggtgcgc agcattgctc tgatcggacc atgggccaat    1320
gccacaaccc aaatgcaagg caactactat ggccctgccc catacctcat cagccctctg    1380
gaagctgcta agaaggccgg ctatcacgtc aactttgaac tcggcacaga gatcgccggc    1440
aacagcacca ctggctttgc caaggccatt gctgccgcca agaagtcgga tgccatcatc    1500
tacctcggtg gaattgacaa caccattgaa caggagggcg ctgaccgcac ggacattgct    1560
tggcccggta atcagctgga tctcatcaag cagctcagcg aggtcggcaa accccttgtc    1620
gtcctgcaaa tgggcggtgg tcaggtagac tcatcctcgc tcaagagcaa caagaaggtc    1680
aactccctcg tctggggcgg atatcccggc cagtcgggag gcgttgccct cttcgacatt    1740
ctctctggca agcgtgctcc tgccggccga ctggtcacca ctcagtaccc ggctgagtat    1800
gttcaccaat tcccccagaa tgacatgaac ctccgacccg atggaaagtc aaaccctgga    1860
cagacttaca tctggtacac cggcaaaccc gtctacgagt ttggcagtgg tctcttctac    1920
accaccttca aggagactct cgccagccac cccaagagcc tcaagttcaa cacctcatcg    1980
atcctctctg ctcctcaccc cggatacact tacagcgagc agattcccgt cttcaccttc    2040
gaggccaaca tcaagaactc gggcaagacg gagtccccat atacggccat gctgtttgtt    2100
cgcacaagca acgctggccc agccccgtac ccgaacaagt ggctcgtcgg attcgaccga    2160
cttgccgaca tcaagcctgg tcactcttcc aagctcagca tccccatccc tgtcagtgct    2220
ctcgcccgtg ttgattctca cggaaaccgg attgtatacc ccggcaagta tgagctagcc    2280
ttgaacaccg acgagtctgt gaagcttgag tttgagttgg tgggagaaga ggtaacgatt    2340
gagaactggc cgttggagga gcaacagatc aaggatgcta cacctgacgc ataa          2394
```

FIGURE 20A

SEQ ID NO: 4
Protein sequence of BxI1, a GH3 family β-xylosidase from *Trichoderma reesei*

mvnnaallaalsallptalaqnnqtyanysaqgqpdlypetlatltlsfpdcehgplknnlvcdssag
yveraqalislftleelilntqn**sgpgvprlglpnyqvwnealhgldranfatkggqfewatsfpmpi**
lttaalnrtlihqiadiistqarafsnsgrygldvyapnvngfrsplwgrgqetpgedafflssayty
eyitgiqggvdpehlkvaatvkhfagydlenwnnqsrlgfdaiitqqdlseyytpqflaaaryaksrs
lmcaynsvngvpscansfflqtllreswgfpewgyvssdcdavynvfnphdyasnqssaaasslragt
didcgqtypwhlnesfvagevsrgeiersvtrlyanlvrlgyfdkknqyrslgwkdvvktdawnisye
aavegivllknd**gtlplskkvrsialigpwanattqmqgnyygpapylispleaakkagyhvnfelgt**
eiagnsttgfakaiaaakksdaiiylggidntieqegadrtdiawpgnqldlikqlsevgkplvvlqm
gggqvdssslksnkkvnslvwggypgqsggvalfdilsgkrapagrlvttqypaeyvhqfpqndmnlr
pdgksnpgqtyiwytgkpvyefgsglfyttfketlashpkslkfntssilsaphpgytyseqipvftf
eaniknsgktespytamlfvrtsnagpapypnkwlvgfdrladikpghssklsipipvsalarvdshg
nrivypgkyelalntdesvklefelvgeevtienwpleeqqikdatpda

FIGURE 20B

SEQ ID NO:5
Nucleotide sequence for Fv3A, a GH3 family enzyme from *Fusarium verticillioides*

atgctgctcaatcttcaggtcgctgccagcgctttgtcgctttctcttttaggtggattggctgaggc
tgctacgccatataccttccggactgtaccaaaggacctttgagcaagaatggaatctgcgatactt
cgttatctccagctaaaagagcggctgctctagttgctgctctgacgcccgaagagaaggtgggcaat
ctggtcaggtaaaatatacccccccccataatcactattcggagattggagctgacttaacgcagcaa
tgcaactggtgcaccaagaatcggacttccaaggtacaactggtggaacgaagcccttcatggcctcg
ctggatctccaggtggtcgctttgccgacactcctccctacgacgcggccacatcatttcccatgcct
cttctcatggccgctgctttcgacgatgatctgatccacgatatcggcaacgtcgtcggcaccgaagc
gcgtgcgttcactaacggcggttggcgcggagtcgacttctggacacccaacgtcaaccctttaaag
atcctcgctggggtcgtggctccgaaactccaggtgaagatgcccttcatgtcagccggtatgctcgc
tatatcgtcaggggtctcgaaggcgataaggagcaacgacgtattgttgctacctgcaagcactatgc
tggaaacgactttgaggactggggaggcttcacgcgtcacgactttgatgccaagattactcctcagg
acttggctgagtactacgtcaggcctttccaggagtgcacccgtgatgcaaaggttggttccatcatg
tgcgcctacaatgccgtgaacggcattcccgcatgcgcaaactcgtatctgcaggagacgatcctcag
agggcactggaactggacgcgcgataacaactggatcactagtgattgtggcgccatgcaggatatct
ggcagaatcacaagtatgtcaagaccaacgctgaaggtgcccaggtagcttttgagaacggcatggat
tctagctgcgagtatactactaccagcgatgtctccgattcgtacaagcaaggcctcttgactgagaa
gctcatggatcgttcgttgaagcgccttttcgaagggcttgttcatactggtttctttgacggtgcca
aagcgcaatggaactcgctcagttttgcggatgtcaacaccaaggaagctcaggatcttgcactcaga
tctgctgtggagggtgctgttcttcttaagaatgacggcactttgcctctgaagctcaagaagaagga
tagtgttgcaatgatcggattctgggccaacgatacttccaagctgcagggtggttacagtggacgtg
ctccgttcctccacagcccgctttatgcagctgagaagcttggtcttgacaccaacgtggcttggggt
ccgacactgcagaacagctcatctcatgataactggaccaccaatgctgttgctgcggcgaagaagtc
tgattacattctctactttggtggtcttgacgcctctgctgctggcgaggacagagatcgtgagaacc
ttgactggcctgagagccagctgacccttcttcagaagctctctagtctcggcaagccactggttgtt
atccagcttggtgatcaagtcgatgacaccgctcttttgaagaacaagaagattaacagtattctttg
ggtcaattaccctggtcaggatggcggcactgcagtcatggacctgctcactggacgaaagagtcctg
ctggccgactacccgtcacgcaatatcccagtaaatacactgagcagattggcatgactgacatggac
ctcagacctaccaagtcgttgccagggagaacttatcgctggtactcaactccagttcttccctacgg
ctttggcctccactacaccaagttccaagccaagttcaagtccaacaagttgacgtttgacatccaga
agcttctcaagggctgcagtgctcaatactccgatacttgcgcgctgccccccatccaagttagtgtc
aagaacaccggccgcattacctccgactttgtctctctggtctttatcaagagtgaagttggacctaa
gccttaccctctcaagacccttgcggcttatggtcgcttgcatgatgtcgcgccttcatcgacgaagg
atatctcactggagtggacgttggataacattgcgcgacggggagagaatggtgatttggttgtttat
cctgggacttacactctgttgctggatgagcctacgcaagccaagatccaggttacgctgactggaaa
gaaggctattttggataagtggcctcaagaccccaagtctgcgtaa

FIGURE 21A

SEQ ID NO:6
Protein sequence of Fv3A

mllnlqvaasalslsllgglaeaatpytlpdctkgplskngicdtslspakraaalvaaltpeekvgn
lvsn**ATGAPRIGLPRYNWWNEALHGLAGSPGGRFADTPPYDAATSFPMPLLMAAAFDDDLIHDIGNVV**
GTEARAFTNGGWRGVDFWTPNVNPFKDPRWGRGSETPGEDALHVSRYARYIVRGLEGDKEQRRIVATC
KHYAGNDFEDWGGFTRHDFDAKITPQDLAEYYVRPFQECTRDAKVGSIMCAYNAVNGIPACANSYLQE
TILRGHWNWTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMDssceytttsdvsdsykqgl
lteklmdrslkrlfeglvhtgffdgakaqwnslsfadvntkeaqdlalrsaveg**AVLLKNDGTLPLKL**
KKKDSVAMIGFWANDTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVAWGPTLQNSSSHDNWTTNAVAA
AKKSDYILYFGGLDASAAGEDRDRENLDWPESQLTLLQKLSSLGKPLVVIQLGDQVDDTALLKNKKIN
SILWVNYPGQDGGTAVMDLLTGRKSPAGRLPVTQYPSKYTEQIGMTDMDLRPTKSLPGRTYRWYSTPV
LPYGFGLHYTkfqakfksnkltfdiqkllkgcsaqysdtcalppiqvsvkntgritsdfvslvfikse
vgpkpyplktlaaygrlhdvapsstkdislewtldniarrgengdlvvypgtytllldeptqakiqvt
ltgkkaildkwpqdpksa

FIGURE 21B

SEQ ID NO:7
Nucleotide sequence for Pf43A, a GH43 family enzyme from *Penicillium funiculosum*

atgcttcagcgatttgcttatattttaccactggctctattgagtgttggagtgaaagccgacaaccc
ctttgtgcagagcatctacaccgctgatccggcaccgatggtatacaatgaccgcgtttatgtcttca
tggaccatgacaacaccggagctacctactacaacatgacagactggcatctgttctcgtcagcagat
atggcgaattggcaagatcatggcattccaatgagcctggccaatttcacctgggccaacgcgaatgc
gtgggccccgcaagtcatccctcgcaacggccaattctacttttatgctcctgtccgacacaacgatg
gttctatggctatcggtgtgggagtgagcagcaccatcacaggtccataccatgatgctatcggcaaa
ccgctagtagagaacaacgagattgatcccaccgtgttcatcgacgatgacggtcaggcatacctgta
ctggggaaatccagacctgtggtacgtcaaattgaaccaagatatgatatcgtacagcgggagcccta
ctcagattccactcaccacggctggatttggtactcgaacgggcaatgctcaacggccgaccacttttt
gaagaagctccatgggtatacaaacgcaacggcatctactatatcgcctatgcagccgattgttgttc
tgaggatattcgctactccacgggaaccagtgccactggtccgtggacttatcgaggcgtcatcatgc
cgacccaaggtagcagcttcaccaatcacgagggtattatcgacttccagaacaactcctacttttc
tatcacaacggcgctcttcccggcggaggcggctaccaacgatctgtatgtgtggagcaattcaaata
caatgcagatggaaccattccgacgatcgaaatgaccaccgccggtccagctcaaattgggactctca
acccttacgtgcgacaggaagccgaaacggcggcatggtcttcaggcatcactacggaggtttgtagc
gaaggcggaattgacgtcgggtttatcaacaatggcgattacatcaaagttaaaggcgtagctttcgg
ttcaggagcccattctttctcagcgcgggttgcttctgcaaatagcggcggcactattgcaatacacc
tcggaagcacaactggtacgctcgtgggcacttgtactgtccccagcactggcggttggcagacttgg
actaccgttacctgttctgtcagtggcgcatctgggacccaggatgtgtattttgttttcggtggtag
cggaacaggatacctgttcaactttgattattggcagttcgcataa

FIGURE 22A

SEQ ID NO:8
Protein sequence of Pf43A

mlqrfayilplallsvgvka**dnpfvqsiytadpapmvyndrvyvfmdhdntgatyynmtdwhlfssad**
manwqdhgipmslanftwananawapqviprngqfyfyapvrhndgsmaigvgvsstitgpyhdaigk
plvenneidptvfidddgqaylywgnpdlwyvklnqdmisysgsptqiplttagfgtrtgnaqrpttf
eeapwvykrngiyyiayaadccsedirystgtsatgpwtyrgvimptqgssftnhegiidfqnnsyff
yhngalpggggyqrsvcveqfkynadgti*ptiemttagpaqigtlnpyvrq*EAETAAWSSGITTEVCS
EGGIDVGFINNGDYIKVKGVAFGSGAHSFSARVASANSGGTIAIHLGSTTGTLVGTCTVPSTGGWQTW
TTVTCSVSGASGTQDVYFVFGGSGTGYLFNFDYWQFa

FIGURE 22B

SEQ ID NO:9
Nucleotide sequence for Fv43E, a GH43 family enzyme from *Fusarium verticillioides*

atgaaggtatactggctcgtggcgtgggccacttctttgacgccggcactggctggcttgattggaca
ccgtcgcgccaccaccttcaacaatcctatcatctactcagactttccagataacgatgtattcctcg
gtccagataactactactacttctctgcttccaacttccacttcagcccaggagcacccgttttgaag
tctaaagatctgctaaactgggatctcatcggccattcaattccccgcctgaactttggcgacggcta
tgatcttcctcctggctcacgttattaccgtggaggtacttgggcatcatccctcagatacagaaaga
gcaatggacagtggtactggatcggctgcatcaacttctggcagacctgggtatacactgcctcatcg
ccggaaggtccatggtacaacaagggaaacttcggtgataacaattgctactacgacaatggcatact
gatcgatgacgatgataccatgtatgtcgtatacggttccggtgaggtcaaagtatctcaactatctc
aggacggattcagccaggtcaaatctcaggtagttttcaagaacactgatattggggtccaagacttg
gagggtaaccgcatgtacaagatcaacgggctctactatatcctaaacgatagcccaagtggcagtca
gacctggatttggaagtcgaaatcaccctggggcccttatgagtctaaggtcctcgccgacaaagtca
ccccgcctatctctggtggtaactcgccgcatcagggtagtctcataaagactcccaatggtggctgg
tacttcatgtcattcacttgggcctatcctgccggccgtcttccggttcttgcaccgattacgtgggg
tagcgatggtttccccattcttgtcaagggtgctaatggcggatggggatcatcttacccaacacttc
ctggcacggatggtgtgacaaagaattggacaaggactgataccttccgcggaacctcacttgctccg
tcctgggagtggaaccataatccggacgtcaactccttcactgtcaacaacggcctgactctccgcac
tgctagcattacgaaggatatttaccaggcgaggaacacgctatctcaccgaactcatggtgatcatc
caacaggaatagtgaagattgatttctctccgatgaaggacggcgaccgggccgggctttcagcgttt
cgagaccaaagtgcatacatcggtattcatcgagataacggaaagttcacaatcgctacgaagcatgg
gatgaatatggatgagtggaacggaacaacaacagacctgggacaaataaaagccacagctaatgtgc
cttctggaaggaccaagatctggctgagacttcaacttgataccaacccagcaggaactggcaacact
atcttttcttacagttgggatggagtcaagtatgaaacactgggtcccaacttcaaactgtacaatgg
ttgggcattctttattgcttaccgattcggcatcttcaacttcgccgagacggctttaggaggctcga
tcaaggttgagtctttcacagctgcatag

FIGURE 23A

SEQ ID NO:10
Protein sequence of Fv43E

mkvywlvawatsltpalaglighrrattfnnpiiysdfpdndvflgpdnyyyfsasnfhfspgapvlk
skdllnwdlighsiprlnfgdgydlppgsryyrggtwasslryrksngqwywigcinfwqtwvytass
pegpwynkgnfgdnncyydngilidddtmyvvygsgevkvsqlsqdgfsqvksqvvfkntdigvqdl
egnrmykinglyyilndspsgsqtwiwkskspwgpyeskvladkvtppisggnsphqgsliktpnggw
yfmsftwaypagrlpvlapitwgsdgfpilvkganggwgssyptlpgtdgvtknwtrtdtfrgtslap
swewnhnpdvnsftvnngltlrtasitkdiyqarntlshrthgdhptgivkidfspmkdgdraglsaf
rdqsayigihrdngkftiatkhgmnmdewngtttdlgqikatanvpsgrtkiwlrlqldtnpagtgnt
ifsyswdgvkyetlgpnfklyngwaffiayrfgifnfaetalggsikvesftaa

FIGURE 23B

SEQ ID NO:11
Nucleotide sequence for Fv43B, a GH43 family enzyme from *Fusarium verticillioides*

atgcgcttctcttggctattgtgccccttctagcgatgggaagtgctcttcctgaaacgaagacgga
tgtttcgacatacaccaaccctgtccttccaggatggcactcggatccatcgtgtatccagaaagatg
gcctctttctctgcgtcacttcaacattcatctccttcccaggtcttcccgtctatgcctcaagggat
ctagtcaactggcgtctcatcagccatgtctggaaccgcgagaaacagttgcctggcattagctggaa
gacggcaggacagcaacagggaatgtatgcaccaaccattcgataccacaagggaacatactacgtca
tctgcgaatacctgggcgttggagatattattggtgtcatcttcaagaccaccaatccgtgggacgag
agtagctggagtgaccctgttaccttcaagccaaatcacatcgaccccgatctgttctgggatgatga
cggaaaggtttattgtgctacccatggcatcactctgcaggagattgatttggaaactggagagctta
gcccggagcttaatatctggaacggcacaggaggtgtatggcctgagggtccccatatctacaagcgc
gacggttactactatctcatgattgccgagggtggaactgccgaagaccacgctatcacaatcgctcg
ggcccgcaagatcaccggcccctatgaagcctacaataacaacccaatcttgaccaaccgcgggacat
ctgagtacttccagactgtcggtcacggtgatctgttccaagataccaagggcaactggtggggtctt
tgtcttgctactcgcatcacagcacagggagtttcacccatgggccgtgaagctgttttgttcaatgg
cacatggaacaagggcgaatggcccaagttgcaaccagtacgaggtcgcatgcctggaaacctcctcc
caaagccgacgcgaaacgttcccggagatgggcccttcaacgctgacccagacaactacaacttgaag
aagactaagaagatccctcctcactttgtgcaccatagagtcccaagagacggtgccttctctttgtc
ttccaagggtctgcacatcgtgcctagtcgaaacaacgttaccggtagtgtgttgccaggagatgaga
ttgagctatcaggacagcgaggtctagctttcatcggacgccgccaaactcacactctgttcaaatat
agtgttgatatcgacttcaagcccaagtccgatgatcaggaagctggaatcaccgttttccgcacgca
gttcgaccatatcgatcttggcattgttcgtcttcctacaaaccaaggcagcaacaagaaatctaagc
ttgccttccgattccgggccacaggagctcagaatgttcctgcaccgaaggtagtaccggtccccgat
ggctgggagaagggcgtaatcagtctacatatcgaggcagccaacgcgacgcactacaaccttggagc
ttcgagccacagaggcaagactctcgacatcgcgacagcatcagcaagtcttgtgagtggaggcacgg
gttcatttgttggtagtttgcttggaccttatgctacctgcaacggcaaaggatctggagtggaatgt
cccaagggaggtgatgtctatgtgacccaatggacttataagcccgtggcacaagagattgatcatgg
tgtttttgtgaaatcagaattgtag

FIGURE 24A

SEQ ID NO:12
Protein sequence of Fv43B

mrfswllcpllamgsalpetktdvst**ytnpvlpgwhsdpsciqkdglflcvtstfisfpglpvyasrd**
lvnwrlishvwnrekqlpgiswktagqqqgmyaptiryhkgtyyviceylgvgdiigvifkttnpwde
sswsdpvtfkpnhidpdlfwdddgkvycathgitlqeidletgelspelniwngtggvwpegphiykr
dgyyylmiaeggtaedhaitiararkitgpyeaynnnpiltnrgtseyfqtvghgdlfqdtkgnwwgl
clatritaqgvspmgreavlfngtwnkgewpklqpvrgrmpgnllpkptrnvpgdgpfnadpdnynlk
ktkkipphfvhhrvprdgafslsskglhivpsrnnvtgsvlpgdeielsgqrglafigrrqthtlfky
svdidfkpksddqeagitvfrtqfdhidlgivrlptnqgsnkksklafrfratgaqnvpapkvvpvpd
gwekgvislhieaanathynlgasshrgktldiatasaslvsggtgsfvgsllgpyatcngkgsgvec
pkggdvyvtqwtykpvaqeidhgvfvksel

FIGURE 24B

SEQ ID NO:13
Nucleotide sequence for Af43A, a GH43 family enzyme from *Aspergillus fumigatus*

```
atggcagctccaagtttatcctaccccacaggtatccaatcgtataccaatcctctcttccctggttg
gcactccgatcccagctgtgcctacgtagcggagcaagacacctttttctgcgtgacgtccactttca
ttgccttccccggtcttcctctttatgcaagccgagatctgcagaactggaaactggcaagcaatatt
ttcaatcggcccagccagatccctgatcttcgcgtcacggatggacagcagtcgggtatctatgcgcc
cactctgcgctatcatgagggccagttctacttgatcgtttcgtacctgggcccgcagactaagggct
tgctgttcacctcgtctgatccgtacgacgatgccgcgtggagcgatccgctcgaattcgcggtacat
ggcatcgacccggatatcttctgggatcacgacgggacggtctatgtcacgtccgccgaggaccagat
gattaagcagtacacactcgatctgaagacgggggcgattggcccggttgactacctctggaacggca
ccggaggagtctggcccgagggcccgcacatttacaagagagacggatactactacctcatgatcgca
gagggaggtaccgagctcggccactcggagaccatggcgcgatctagaacccggacaggtccctggga
gccatacccgcacaatccgctcttgtcgaacaagggcacctcggagtacttccagactgtgggccatg
cggacttgttccaggatgggaacggcaactggtgggccgtggcgttgagcacccgatcagggcctgca
tggaagaactatcccatgggtcgggagacggtgctcgcccccgccgcttgggagaagggtgagtggcc
tgtcattcagcctgtgagaggccaaatgcaggggccgtttccaccaccaaataagcgagttcctcgcg
gcgagggcggatggatcaagcaacccgacaaagtggatttcaggcccggatcgaagataccggcgcac
ttccagtactggcgatatcccaagacagaggattttaccgtctcccctcggggccacccgaatactct
tcggctcacaccctccttttacaacctcaccggaactgcggacttcaagccggatgatggcctgtcgc
ttgttatgcgcaaacagaccgacaccttgttcacgtacactgtggacgtgtcttttgaccccaaggtt
gccgatgaagaggcgggtgtgactgttttccttacccagcagcagcacatcgatcttggtattgtcct
tctccagacaaccgaggggctgtcgttgtccttccggttccgcgtggaaggccgcggtaactacgaag
gtcctcttccagaagccaccgtgcctgttcccaaggaatggtgtggacagaccatccggcttgagatt
caggccgtgagtgacaccgagtatgtctttgcggctgccccggctcggcaccctgcacagaggcaaat
catcagccgcgccaactcgttgattgtcagtggtgatacgggacggtttactggctcgcttgttggcg
tgtatgccacgtcgaacgggggtgccggatccacgcccgcatatatcagcagatggagatacgaagga
cggggccagatgattgattttggtcgagtggtcccgagctactga
```

FIGURE 25A

SEQ ID NO:14
Protein sequence of Af43A

maapslsyptgiqs**ytnplfpgwhsdpscayvaeqdtffcvtstfiafpglplyasrdlqnwklasni**
fnrpsqipdlrvtdgqqsgiyaptlryhegqfylivsylgpqtkgllftssdpyddaawsdplefavh
gidpdifwdhdgtvyvtsaedqmikqytldlktgaigpvdylwngtggvwpegphiykrdgyyylmia
eggtelghsetmarsrtrtgpwepyphnpllsnkgtseyfqtvghadlfqdgngnwwavalstrsgpa
wknypmgretvlapaawekgewpviqpvrgqmqgpfpppnkrvprgeggwikqpdkvdfrpgskipah
fqywrypktedftvsprghpntlrltpsfynltgtadfkpddglslvmrkqtdtlftytvdvsfdpkv
adeeagvtvfltqqqhidlgivllqtteglslsfrfrvegrgnyegplpeatvpvpkewcgqtirlei
qavsdteyvfaaaparhpaqrqiisranslivsgdtgrftgslvgvyatsnggagstpayisrwryeg
rgqmidfgrvvpsy

FIGURE 25B

SEQ ID NO:15
Nucleotide sequence of Fv51A, a GH51 family enzyme from *Fusarium verticillioides*

atggttcgcttcagttcaatcctagcggctgcggcttgcttcgtggctgttgagtcagtcaacatcaaggtcgac
agcaagggcggaaacgctactagcggtcaccaatatggcttccttcacgaggttggtattgacacaccactggcg
atgattgggatgctaacttggagctaggatatcaacaattccggtgatggtggcatctacgctgagctcatccgc
aatcgtgctttccagtacagcaagaaataccctgtttctctatctggctggagacccatcaacgatgctaagctc
tccctcaaccgtctcgacactcctctctccgacgctctccccgtttccatgaacgtgaagcctggaaagggcaag
gccaaggagattggtttcctcaacgagggttactggggaatggatgtcaagaagcaaaagtacactggctctttc
tgggttaagggcgcttacaagggccactttacagcttctttgcgatctaaccttaccgacgatgtctttggcagc
gtcaaggtcaagtccaaggccaacaagaagcagtgggttgagcatgagtttgtgcttactcctaacaagaatgcc
cctaacagcaacaacacttttgctatcacctacgatcccaaggtgagtaacaatcaaaactgggacgtgatgtat
actgacaatttgtagggcgctgatggagctcttgacttcaacctcattagcttgttccctcccacctacaagggc
cgcaagaacggtcttcgagttgatcttgccgaggctctcgaaggtctccacccccgtaaggtttaccgtctcacgt
gtatcgtgaacagtcgctgacttgtagaaaagagcctgctgcgcttccccggtggtaacatgctcgagggcaaca
ccaacaagacctggtgggactggaaggataccctcggacctctccgcaaccgtcctggtttcgagggtgtctgga
actaccagcagacccatggtcttggaatcttggagtacctccagtgggctgaggacatgaaccttgaaatcagta
ggttctataaaattcagtgacggttatgtgcatgctaacagatttcagttgtcggtgtctacgctggcctctccc
tcgacggctccgtcacccccaaggaccaactccagcccctcatcgacgacgcgctcgacgagatcgaattcatcc
gaggtcccgtcacttcaaagtggggaaagaagcgcgctgagctcggccaccccaagcctttcagactctcctacg
ttgaagtcggaaacgaggactggctcgctggttatcccactggctggaactcttacaaggagtaccgcttcccca
tgttcctcgaggctatcaagaaagctcaccccgatctcaccgtcatctcctctggtgcttctattgaccccgttg
gtaagaaggatgctggtttcgatattcctgctcctggaatcggtgactaccacccttaccgcgagcctgatgttc
ttgttgaggagttcaacctgtttgataacaataagtatggtcacatcattggtgaggttgcttctacccacccca
acggtggaactggctggagtggtaaccttatgccttacccctggtggatctctggtgttggcgaggccgtcgctc
tctgcggttatgagcgcaacgccgatcgtattcccggaacattctacgctcctatcctcaagaacgagaaccgtt
ggcagtgggctatcaccatgatccaattcgccgccgactccgccatgaccacccgctccaccagctggtatgtct
ggtcactcttcgcaggccaccccatgacccatactctccccaccaccgccgacttcgacccccctctactacgtcg
ctggtaagaacgaggacaagggaactcttatctggaagggtgctgcgtataacaccaccaagggtgctgacgttc
ccgtgtctctgtccttcaagggtgtcaagcccggtgctcaagctgagcttactcttctgaccaacaaggagaagg
atccttttgcgttcaatgatcctcacaagggcaacaatgttgttgatactaagaagactgttctcaaggccgatg
gaaagggtgctttcaacttcaagcttcctaacctgagcgtcgctgttcttgagaccctcaagaagggaaagccctt
actctagctag

FIGURE 26A

SEQ ID NO:16
Protein sequence of Fv51A

mvrfssilaaaacfvavesvnikvdskggnatsghqygflhedinnsgdggiyaelirnrafqyskky
pvslsgwrpindaklslnrldtplsdalpvsmnvkpgkgkakeigflnegywgmdvkkqkytgsfwvk
gaykghftaslrsnltddvfgsvkvkskankkqwvehefvltpnknapnsnntfaitydpkgadgald
fnlislfpptykgrknglrvdlaealeglhpsllrfpggnmlegntnktwwdwkdtlgplrnrpgfeg
vwnyqqthglgileylqwaedmnleiivgvyaglsldgsvtpkdqlqpliddaldeiefirgpvtskw
gkkraelghpkpfrlsyvevgnedwlagyptgwnsykeyrfpmfleaikkahpdltvissgasidpvg
kkdagfdipapgigdyhpyrepdvlveefnlfdnnkyghii**gevasthpnggtgwsgnlmpypwwisg**
vgeavalcgyernadripgtfyapilknenrwqwaitmiqfaadsamttrstswyvwslfaghpmtht
lpttadfdplyyvagknedkgtliwkgaaynttkgadvpvslsfkgvkpgaqaeltlltnkekdpfaf
ndphkgnnvvdtkktvlkadgkgafnfklpnlsvavletlkkgkpyss

FIGURE 26B

SEQ ID NO:17
Nucleotide sequence for Xyn3, a GH10 xylanase from *Trichoderma reesei*

atgaaagcaaacgtcatcttgtgcctcctggccccctggtcgccgctctccccaccgaaaccatcca
cctcgaccccgagctcgccgctctccgcgccaacctcaccgagcgaacagccgacctctgggaccgcc
aagcctctcaaagcatcgaccagctcatcaagagaaaaggcaagctctactttggcaccgccaccgac
cgcggcctcctccaacgggaaaagaacgcggccatcatccaggcagacctcggccaggtgacgccgga
gaacagcatgaagtggcagtcgctcgagaacaaccaaggccagctgaactggggagacgccgactatc
tcgtcaactttgcccagcaaaacggcaagtcgatacgcggccacactctgatctggcactcgcagctg
cctgcgtgggtgaacaatatcaacaacgcggatactctgcggcaagtcatccgcacccatgtctctac
tgtggttgggcggtacaagggcaagattcgtgcttgggtgagttttgaacaccacatgccccttttct
tagtccgctcctcctcctcttggaacttctcacagttatagccgtatacaacattcgacaggaaattt
aggatgacaactactgactgacttgtgtgtgtgatggcgataggacgtggtcaatgaaatcttcaacg
aggatggaacgctgcgctcttcagtcttttccaggctcctcggcgaggagtttgtctcgattgccttt
cgtgctgctcgagatgctgacccttctgcccgtctttacatcaacgactacaatctcgaccgcgccaa
ctatggcaaggtcaacgggttgaagacttacgtctccaagtggatctctcaaggagttcccattgacg
gtattggtgagccacgacccctaaatgtcccccattagagtctctttctagagccaaggcttgaagcc
attcagggactgacacgagagccttctctacaggaagccagtcccatctcagcggcggcggaggctct
ggtacgctgggtgcgctccagcagctggcaacggtacccgtcaccgagctggccattaccgagctgga
cattcagggggcaccgacgacggattacacccaagttgttcaagcatgcctgagcgtctccaagtgcg
tcggcatcaccgtgtggggcatcagtgacaaggtaagttgcttcccctgtctgtgcttatcaactgta
agcagcaacaactgatgctgtctgtctttacctaggactcgtggcgtgccagcaccaaccctcttctg
tttgacgcaaacttcaaccccaagccggcatataacagcattgttggcatcttacaatag

FIGURE 27A

SEQ ID NO:18
Protein sequence for Xyn3

mkanvilcllaplvaalptetihldpelaalranltertadlwdrqasq**sidqlikrkgklyfgtatd**
rgllqreknaaiiqadlgqvtpensmkwqslennqgqlnwgdadylvnfaqqngksirghtliwhsql
pawvnninnadtlrqvirthvstvvgrykgkirawdvvneifnedgtlrssvfsrllgeefvsiafra
ardadpsarlyindynldranygkvnglktyvskwisqgvpidgigsqshlsggggsgtlgalqqlat
vpvtelaiteldiqgapttdytqvvqaclsvskcvgitvwgisdkdswrastnpllfdanfnpkpayn
sivgilq

FIGURE 27B

SEQ ID NO:19
Nucleotide sequence xlnA, a xylanase from *Aspergillus tubingensis*

```
aacgtctgca gtcccgtact gttttaccaaa atgccaggcc actggtggat aaacaacttt      60
gtaatacgtt gccggagtca gcccctactc cctgaagggt tccccactcc tagttactta      120
ctactgggta gtaggctcct agaggtgggt aaaggttttag ccccagaccta      180
gttttatgttt ggctaggcag gacctgggta agttgatggc tcctgcattc ctaacctgagt      240
attttccagct ataagcgaga tttgccatac actttcagcga gtccggatgg tccgcgccga      300
ggttgaccct gccttcatca cctacacaaa gaactcctcg gccaactccc ggtggccttc      360
gagcttccaaa gtaccttcgc gaccttttgc cagtgattct cgcagcgttt actgagccta      420
aggcttgata caataaataa agagacataa ccttgcagta catacgtctt gtatgagcga      480
ggaactgtgt tcagtagtag atcagtgggt acattaatcat gaacatgact tctgagccag      540
aaaaccttct gcagggaacc ggtgaagaaa ccccacttcc ccgcctccac taactgcagc      600
ccttttatcc gcctgccgtc catttagcca aatgtagtcc attttagccaa gtgcggtcca      660
ttttagccaag tccagtgctt aggttggtgg ctacacagga aacggccatg aatgtagaca      720
caactataga actgtcccta gaaaataggct cgaggttgtt agagcgttta aggtgatgcg      780
gcaaaatgca tatgactgag ttgcttcaac gtgcagggga aagggataaa tagtctttt      840
cgcagaatat aaatagaggt agagcggggt cgcagcaata ttgacccagga cagggcttct      900
tttccagttg catacatcca ttcacagcat tcagctttct tcaatcatca tgaaggtcac      960
tgcggctttt gcaggtcttt tggtcacggc attcgccgct cctgccccag aaacttgatct     1020
ggtgtcgtga agtgccgtta tcaactacgt gcaaaactac aacggcaacc tcggtgattt     1080
cacctacgac gagagtgccg gaacatttte catgtactgg gaagatggag tgagctccga     1140
ctttgtcgtt ggtctcgggct ggatcgactg gtcgttctaag tgagtgactg tatttctttaa     1200
ccaaggtcta ggatctaacg tctttcagcg ctatcacctа ctctgccgaa tacagcgctt     1260
ctggctcggc ttcctacctc gctgtgtacg gctgggtcaa ctatcctcaa gctgagtact     1320
acatcgtcga ggattacggt gattataacc cttgcagttc ggccacaagc cctggtaccg     1380
tgtactctga tggaagcacc taccaagtct gcaccgacac tcgaacaaac gaaccgtcca     1440
tcacgggaac aagcacgttc acgcagtact tctccgttcg agagagcacg cgcacatctg     1500
gaacggtcac tgttgccaac catttcaact tctgggcgca ccatgggttc ggcaatagcg     1560
acttcaatta tcaggtcgtg gcggtggaag catggagcgg tgctggcagc gctagtgtca     1620
caatcttctc ttgagacgatt agtgccctag tagtcggaag atatcaacgc ggcagttgc     1680
tctcaggtgg tgtgatgatc ggatccggtc tctgggggtta cattgaggct gtataagttg     1740
ttgtgggccc gagctgtcag cggctgcgtt ttcagcttgc acagataatc aactctcgtt     1800
ttctatctct tgcgtttcct cgctgcttat cctatccata gataattatt ttgcccacta     1860
ccacaacttg ttcggtcgca gtagtcactc cgagcaaggc attgggaaat gggggatgcg     1920
gggtgctgcg taccctctaa cctagggcat tttaaaggat atttaccctc cagatattct     1980
atagatacag acttcttagg actgcgggta atatagagag cgaaatttct acagttcgat     2040
gcagttcaat gcga                                                       2054
```

FIGURE 28A

SEQ ID NO:20
Protein sequence for xlnA

MKVTAAFAGLLVTAFAAPAPEPDLVSRSAGINYVQNYNGNLGDFTYDESAGTFSMYWEDGVSSDFVVGLG
WTTGSSNAITYSAEYSASGSASYLAVYGWVNYPQAEYYIVEDYGDYNPCSSATSLGTVYSDGSTYQVCTD
TRTNEPSITGTSTFTQYFSVRESTRTSGTVTVANHFNFWAHHGFGNSDFNYQVVAVEAWSGAGSASVTIS
S

FIGURE 28B

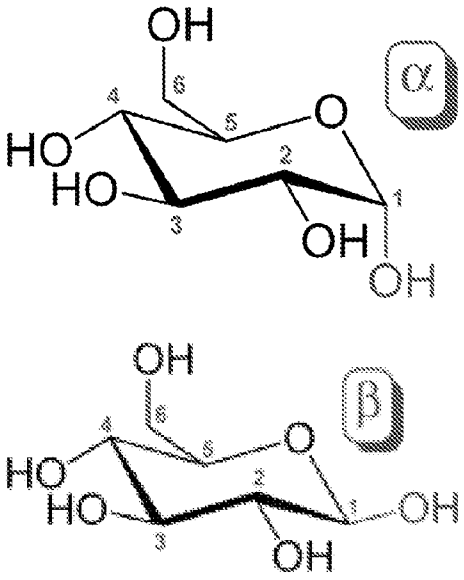


FIGURE 29

SEQ ID NO:21
Nucleotide sequence for Gz43A, a GH43 family enzyme from *Gibberella zeae*

```
atgaagtccaagttgttattcccactcctctctttcgttggtcaaagccttgccaccaacgacgactgtcctccccatcactag
tagatcgactgcggatccttcggctcatgtccttaacgacaccttgtggctctacccgtctcatgacatcgatgctgcatttg
agaatcatcctgatgcaggccagtacgccatgagagattaccatgtccactctatcgacaagatctacggttccctgccggtc
gatcacggtacggccctgtcagtggaggatgcccctgggcctctcgacagatgtgggctcctgacgctgcccacaacaacgg
caaatactacctatacttcccgccaaagacaaggatgatatcttcagaatcggcgtcgctgtctcaccaacccccggcggac
cattcgtccccgacaagagttggatccctcacactttcagcatcgaccccgccagtttcgtcgatgatgatgacagacgcctac
ttggcatggggtggtatcatgggtggccagcctcaacgatgccaggacaagaacaagtacaacgaatctggcactgacgccagg
aaacggcaccgctgccttgagccctcagattgccaagctgacgcaaggacatgcacactctggcagagaagcctcgcgacatgc
tcattcttgaccccaagactggcaagccgctcctttctgaggatgaagaccgacgctcttcgaaggaccctcgattcacaag
cgcaacaagatttactacctcacctactctactggcacaacccactacttgtctatgcgacttcaaagacccccctatggtcc
ttacacctaccagggcagaatcctggagccagttgatggctcgactactcactctagcatcgtcaagtaccagggtcagtggt
ggctattttatcacgatgccaagacatctggcaaggactatcttcgccaggtaaaggctaagaagatttggtacgatagcaaa
ggaaacgatcttgacaaagaagccttga
```

FIGURE 30A

SEQ ID NO:22
Protein sequence of Gz43A

mksklfpllsfvgqslatnddcplitsrwtadpsahvfndtlwlypshdid**agfendpdggqyamrdyh**
vysidkiygslpvdhgtalsvedvpwasrqmwapdaahkngkyylyfpakdkddifrigvavsptpgg
pfvpdkswiphtfsidpasfvddddraylawggimggqlqrwqdknkynesgtepgngtaalspqiak
lskdmhtlaekprdmlildpktgkpllsededrrffegpwihkrnkiyyltystgtthylvyatsktp
ygpytyqgrilepvdgwtthssivkyqgqwwlfyhdaktsgkdylrqvkakkiwydskgkiltkkp

FIGURE 30B

SEQ ID NO:23
Nucleotide sequence for Fo43A, a GH43 family enzyme from *Fusarium oxysporum*

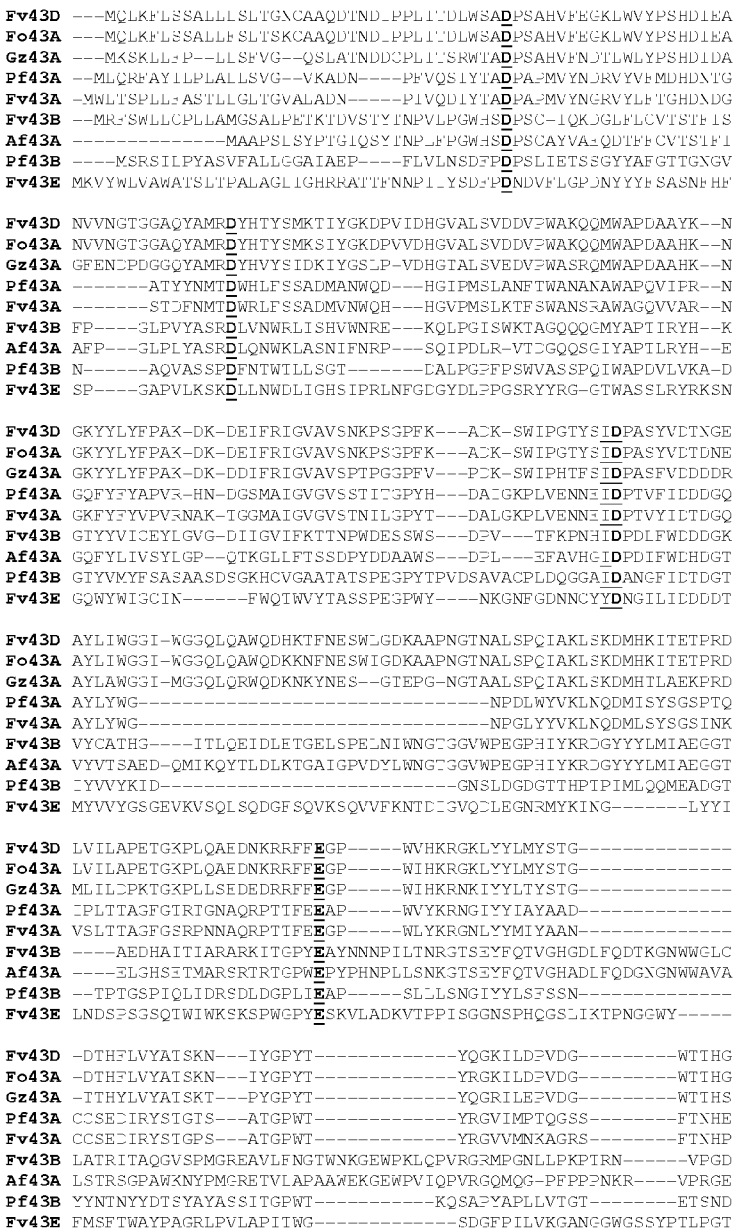
atgcagctcaagtttctgtcttcagcattgctgttctctctgaccagcaaatgcgctgcgcaagacac
taatgacattcctcccctgatcaccgacctctggtccgcagatccctcggctcatgttttcgaaggca
agctctgggtttacccatctcacgacatcgaagccaatgttgtcaacggcacaggaggcgctcaatac
gccatgagggattaccatacctactccatgaagagcatctatggtaaagatcccgttgtcgaccacgg
cgtcgctctctcagtcgatgacgttccctgggcgaagcagcaaatgtgggctcctgacgcagctcata
agaacggcaaatattatctgtacttccccgccaaggacaaggatgagatcttcagaattggagttgct
gtctccaacaagcccagcggtcctttcaaggccgacaagagctggatccctggcacgtacagtatcga
tcctgctagctacgtcgacactgataacgaggcctacctcatctggggcggtatctggggcggccagc
tccaagcctggcaggataaaaagaactttaacgagtcgtggattggagacaaggctgctcctaacggc
accaatgccctatctcctcagatcgccaagctaagcaaggacatgcacaagatcaccgaaacaccccg
cgatctcgtcattctcgcccccgagacaggcaagcctcttcaggctgaggacaacaagcgacgattct
tcgagggcccttggatccacaagcgcggcaagctttactacctcatgtactccaccggtgatacccac
ttccttgtctacgctacttccaagaacatctacggtccttatacctaccggggcaagattcttgatcc
tgttgatgggtggactactcatggaagtattgttgagtataagggacagtggtggcttttctttgctg
atgcgcatacgtctggtaaggattaccttcgacaggtgaaggcgaggaagatctggtatgacaagaac
ggcaagatcttgcttcaccgtccttag

FIGURE 31A

SEQ ID NO:24
Protein sequence of Fo43A

Mqlkflssallfsltskcaaqdtndipplitdlwsadpsahvfegklwvypshdieanvvngtggaqy
amrdyhtysmksiygkdpvvdhgvalsvddvpwakqqm**wapdaahkngkyylyfpakdkdeifrigva**
vsnkpsgpfkadkswipgtysidpasyvdtdneayliwggiwggqlqawqdkknfneswigdkaapng
tnalspqiaklskdmhkitetprdlvilapetgkplqaednkrrffegpwihkrgklyylmystgdth
flvyatskniygpytyrgkildpvdgwtthgsiveykgqwwlffadahtsgkdylrqvkarkiwydkn
gkillhrp

FIGURE 31B


Fv43D ---MQLKFLSSALLLSLTGXCAAQDTND-PPLTTDLWSADPSAHVFEGKLWVYPSHDIEA
Fo43A ---MQLKFLSSALLFSLTSKCAAQDTND-PPLTTDLWSADPSAHVFEGKLWVYPSHDIEA
Gz43A ---MKSKLLFP--LLSFVG--QSLATNDDCPLTTSRWTADPSAHVFNGTLWLYPSHDIDA
Pf43A --MLQRFAYILPLALLSVG--VKADN----PFVQSIYTADPAPMVYNGRVYVFMDHDNTG
Fv43A -MWLTSPLLPASTLLGLTGVALADN-----PIVQDIYTADPAPMVYNGRVYLFTGHDNDG
Fv43B --MRFSWLLCPLLAMGSALPETKTDVSTYINPVLPGWHSDPSC-IQKDGLFLCVTSTFIS
Af43A ---------------MAAPSLSYPTGIQSYINPAFPGWHSDPSCAYVAPQDTFFCVTSTFI
Pf43B ----MSRSILPYASVFALLGGAIAEP----FLVLNSDFPDPSLIETSSGYYAFGTTGNGV
Fv43E MKVYWLVAWATSLTPALAGLIGHRRATTFNNPIAYSDFPDNDVFLGPDNYYYFSASNFHF

Fv43D NVVNGTGGAQYAMRDYHTYSMKTIYGKDPVIDHGVALSVDDVPWAKQQMWAPDAAYK--N
Fo43A NVVNGTGGAQYAMRDYHTYSMKSIYGKDPVVDHGVALSVDDVPWAKQQMWAPDAAHK--N
Gz43A GFENDPDGGQYAMRDYHVYSIDKIYGSDP-VDHGIALSVEDVPWASRQMWAPDAAHK--N
Pf43A -------ATYYNMTDWHLFSSADMANWQD---HGIPMSLANFTWANANAWAPQVIPR--N
Fv43A -------STDFNMTDWRLFSSADMVNWQH---HGVPMSLKTFSWANSRAWAGQVVAR--N
Fv43B FP----GLPVYASRDLVNWRLISHVWNRE---KQLPGISWKTAGQQQGMYAPTIRYH--K
Af43A AFP---GLPLYASRDLQNWKLASNIFNRP---SQIPDLR-VTDGQQSGIYAPTLRYH--E
Pf43B N------AQVASSPDFNTWILLSGT-------DALPGFFPSWVASSPQIWAPDVLVKA-D
Fv43E SP----GAPVLKSKDLLNWDLIGHSIPRLNFGDGYDLPPGSRYYRG-GTWASSLRYRKSN

Fv43D GKYYLYFPAK-DK-DEIFRIGVAVSNKPSGPFK---ADK-SWIPGTYSIDPASYVDTNGE
Fo43A GKYYLYFPAK-DK-DEIFRIGVAVSNKPSGPFK---ADK-SWIPGTYSIDPASYVDTDNE
Gz43A GKYYLYFPAK-DK-DDIFRIGVAVSPTPGGPFV---PDK-SWIPHTFSIDPASFVDDDDR
Pf43A GQFYFYAPVR-HN-DGSMAIGVGVSSTITGPYH---DAIGKPLVENNEIDPTVFIDDDGQ
Fv43A GKFYFYVPVRNAK-IGGMAIGVGVSTNILGPYT---DALGKPLVENNEIDPTVYILTDGQ
Fv43B GTYYVICEYLGVG-DIIGVIFKTTNPWDESSWS---DPV---TFKPNHIDPDLFWDDDGK
Af43A GQFYLIVSYLGP--QTKGLLFTSSDPYDDAAWS---DPL---EFAVHGIDPDIFWLHDGT
Pf43B GTYVMYFSASAASDSGKHCVGAATATSPEGPYTPVDSAVACPLDQGGAIDANGFILTDGT
Fv43E GQWYWIGCIN------FWQIWVYTASSPEGPWY----NKGNFGDNNCYYDNGILILDDDT

Fv43D AYLIWGGI-WGGQLQAWQDHKTFNESWLGDKAAPNGTNALSPQIAKLSKDMHKITETPRD
Fo43A AYLIWGGI-WGGQLQAWQDKKNFNESWIGDKAAPNGTNALSPQIAKLSKDMHKITETPRD
Gz43A AYLAWGGI-MGGQLQRWQDKNKYNES--GIEPG-NGTAALSPQIAKLSKDMHILAEKPRD
Pf43A AYLYWG----------------------------------NPDLWYVKLNQDMISYSGSPTQ
Fv43A AYLYWG----------------------------------NPGLYYVKLNQDMLSYSGSINK
Fv43B VYCATHG----ITLQEIDLETGELSPELNIWNGTGGVWPEGPHIYKRDGYYYLMIAEGGT
Af43A VYVTSAED-QMIKQYTLDLKTGAIGPVDYLWNGTGGVWPEGPHIYKRDGYYYLMIAEGGT
Pf43B IYVVYKID----------------------------GNSLDGDGTTHPIPIMLQQMEADGT
Fv43E MYVVYGSGEVKVSQLSQDGFSQVKSQVVFKNTDIGVQILEGNRMYKING-------LYYI

Fv43D LVILAPETGKPLQAEDNKRRFFEGP-----WVHKRGKLYYLMYSTG----------
Fo43A LVILAPETGKPLQAEDNKRRFFEGP-----WIHKRGKLYYLMYSTG----------
Gz43A MLILDPKTGKPLLSEDEDRRFFEGP-----WIHKRNKIYYLTYSTG----------
Pf43A IPLTTAGFGTRIGNAQRPTIFEEAP-----WVYKRNGIYYIAYAAD----------
Fv43A VSLTTAGFGSRPNNAQRPTIFEEGP-----WLYKRGNLYYMIYAAN----------
Fv43B ----AEDHAITIARARKITGPYEAYNNNPILTNRGTSEYFQTVGHGDLFQDTKGNWWGLC
Af43A ----ELGHSETMARSRTRTGPWEPYPHNPLLSNKGTSEYFQTVGHADLFQDGKGNWWAVA
Pf43B --TPTGSPIQLIDRSDLDGPLIEAP-----SLILSNGIYYLSFSSN----------
Fv43E LNDSPSGSQTWIWKSKSPWGPYESKVLADKVTPPISGGNSPHQGSLIKTPNGGWY-----

Fv43D -DTHFLVYATSKN---IYGPYT----------YQGKILDPVDG---------WTIHG
Fo43A -DTHFLVYATSKN---IYGPYT----------YRGKILDPVDG---------WTIHG
Gz43A -TTHYLVYATSKT---PYGPYT----------YQGRILEPVDG---------WTIHS
Pf43A CCSEDIRYSTGTS---ATGPWT----------YRGVIMPTQGSS--------FTKHE
Fv43A CCSEDIRYSTGPS---ATGPWT----------YRGVVMNKAGRS--------FTKHP
Fv43B LATRITAQGVSPMGREAVLFNGTWNKGEWPKLQPVRGRMPGNLLPKPIRN------VPGD
Af43A LSTRSGPAWKNYPMGRETVLAPAAWEKGEWPVIQPVRGQMQG-PFPPPNKR----VPRGE
Pf43B YYNTNYYDTSYAYASSITGPWT----------KQSAPYAPLLVTGT--------ETSND
Fv43E FMSFTWAYPAGRLPVLAPITWG----------SDGFPILVKGANGGWGSSYPTLPGT

FIGURE 32-1

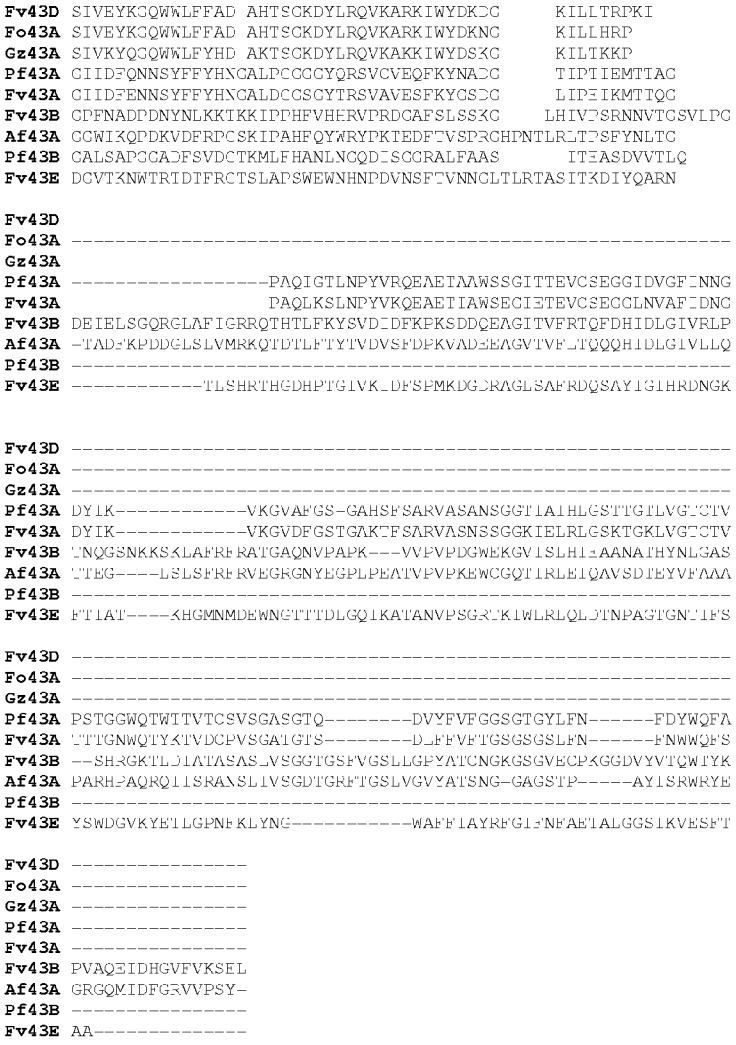


FIGURE 32-2

SEQ ID NO: 25
Protein sequence of XynB3, a GH43 family enzyme from *Geobacillus stearothermophilus*

MSKIKNPILTGFHPDPSICRVGDDYYIAVSTFEWFPGVRIYHSKDLKNWRLVARPLNRLSQLNMIGNPDS
GGVWAPHLSYSDGKFWLIYTDVKVVEGQWKDGHNYLVTCDTIDGAWSDPIYLNSSGFDPSLFHDEDGRKY
LVNMYWDHRVDHHPFYGIVLQEYSVEQKKLVGEPKIIFKGTDLRITEGPHLYKINGYYYLLTAEGGTRYN
HAATIARSTSLYGPYEVHPDNPLLTSWPYPRNPLQKAGHASIVHTHTDEWFLVHLTGRPLPREGQPLLEH
RGYCPLGRETAIQRLEWKDGWPYVVGGNGPSLEIDGPSVEEVSWEKDYDEKDDFDGDTLNHHFQTLRIPL
GEDIATLKARPGHLRLYGRESLTSRFTQAFVARRWQHFHFVAETKVSFRPTTFQQSAGLVNYYNTQNWTT
LQITWHEEKGRILELMTCDHLVVDQPLRGREIVVPDDIEYVYLRVTVQATTYKYSYSFDGMNWIDLPVTF
ESYKLSDDYIKSRAAFTGAFVGMHCRDGSGQNNYADFDYFLYKEL

FIGURE 33

SEQ ID NO:26
Protein sequence of Bgl1, a GH3 family β-glucosidase from *Trichoderma reesei*

mryrtaaalalatgpfaradshstsgasaeavvppagtpwgtaydkakaalaklnlqdkvgivsgv**gw**
nggpcvgntspaskisypslclqdgplgvrystgstaftpgvqaastwdvnlirergqfigeevkasg
ihvilgpvagplgktpqggrnwegfgvdpyltgiamgqtingiqsvgvqatakhyilneqelnretis
snpddrtlhelytwpfadavqanvasvmcsynkvnttwacedqytlqtvlkdqlgfpgyvmtdwnaqh
ttvqsansgldmsmpgtdfngnnrlwgpaltnavnsnqvptsrvddmvtrilaawyltgqdqagypsf
nisrnvqgnhktnvraiardgivllk**ndanilplkkpasiavvgsaaiignharnspscndkgcddga**
lgmgwgsgavnypyfvapydaintrassqgtqvtlsntdntssgasaargkdvaivfitadsgegyit
vegnagdrnnldpwhngnalvqavagansnvivvvhsvgaiileqilalpqvkavvwaglpsqesgna
lvdvlwgdvspsgklvytiakspndyntrivsggsdsfseglfidykhfddanitpryefgyglsytk
fnysrlsvlstaksgpatgavvpggpsdlfqnvatvtvdiansgqvtgaevaqlyitypssaprtppk
qlrgfaklnltpgqsgtatfnirrrdlsywdtasqkwvvpsgsfgisvgassrdirltstlsva

FIGURE 34

SEQ ID NO:39
Nucleotide sequence of XlnD, a GH3 enzyme from *Aspergillus japonicus*

ATGGCTGTGGCGGCTCTTGCTCTGCTGGCTCTACTGCCTCAAGCTCTGGGGCAACATAACAGCAGCTACG
TGGATTACAACGTCGAAGCCAATCCGGACTTGTTTCCACAATGTCTAGACACAATCTCCCTGTCCTTCCC
CGACTGCCAGAGCGGTCCTCTGAGCAAGAACCTCGTCTGCGACTCGACTGCCTCGCCCTATGACCGCGCC
GCGGCTCTGGTCTCCCTCTTCACCCTCGAGGAACTTATCGCCAACACTGGTAACACCAGCCCGGGTGTCC
CTCGTCTGGGTCTGCCTCCATACCAGGTCTGGAGTGAGGCCCTGCATGGCCTGGCTCGCGCCAACTTCAC
CGACAACGGGGCTTACAGCTGGGCGACGTCCTTCCCCTCACCCATTCTCTCCGCAGCGGCCTTCAATCGC
ACCCTGATCAACCAGATCGCCTCCATTATTTCGACTCAGGGCCGTGCCTTCAACAACGCCGGCCGCTTTG
GCCTCGACGTCTACTCGCCAAACATCAATACCTTCCGCCATCCAGTCTGGGGTCGTGGACAGGAAACTCC
TGGCGAGGATGCGTACACTCTTACGGCCGCCTACGCCTACGAATACATCACGGGTATCCAGGGTGGTGTG
AACCCAGAGCATCTGAAGCTCGCCGCGACAGCCAAGCACTTTGCCGGCTACGACATCGAGAACTGGGACA
ACCACTCCCGGCTGGGGAACGATGTCAACATTACGCAGCAAGACCTGGCCGAGTACTACACGCCGCAGTT
CCTCGTCGCCGCGCGCGACGCCCACGTCCACAGCTTCATGTGCTCCTACAACGCCGTTAACGGAGTGCCC
AGCTGCTCCAACACCTTCTTTCTGCAGACCCTCCTGCGCGACACCTTCTCCTTCGTCGACCACGGCTACG
TCTCCGGCGACTGCGGCGCCGTCTACGGCGTCTTCAACCCCCACGGCTACGCGGCCAACGAGCCCAGCGC
CGCCGCCGATGCCATCCTCGCCGGCACTGACATTGACTGCGGCACCTCCTATCAATATCACTTCAACGAG
TCCATCACCACCGGGGCTGTCGCCCGCGACGACATCGAGCGTGGTTTCATCCGGCTGTACGCCAACCTCG
TCGAGCTGGGCTACTTCGACGGCAACAGCAGCAGCAGCAACCCGTACCGCAGCCTGGGCTGGCCCGACGT
CCAGAAGACAGACGCATGGAACATTTCCTACGAGGCGGCAGTCGAAGGCATCGTCCTCCTGAAGAACGAC
GGCACCCTCCCTCTTGCCTCCCCCTCCGAGGGCAAGAACAAATCCATCGCCCTCATCGGCCCCTGGGCCA
ACGCCACCACCCAGCTCCAGGGTAACTACTACGGCGACGCGCCATACCTCATCAGCCCGGTCGACGCCTT
CACGGCCGCCGGGTACACAGTACACTACGCCCCCGGCACGGAGATCTCGACGAACTCGACGGCGAACTTC
AGCGCCGCGCTCTCCGCGGCGCGCGCCGCCGACACCATCGTATTCTTGGGGGGCATCGATAACACCATCG
AAGCCGAAGCCCAAGACCGCAGCTCGATCGCCTGGCCCGGCAACCAACTCGAGCTGATCTCGCAACTCGC
GGCGCAGAAATCCGACGACCAGCCCCTGGTGGTGTACCAGATGGGCGGCGGCCAGGTCGACTCCTCCGCC
CTCAAATCCAACGCGAAGGTCAACGCCCTCCTCTGGGGCGGCTACCCGGGCCAATCCGGCGGCCTCGCCC
TGCGCGACATCCTCACGGGCGCGCGCGCCCCCGCCGGCCGCCTCACCACGACCCAGTACCCGGCCGCCTA
CGCCGAGAGCTTCTCGGCCCTCGACATGAACCTGCGGCCGAATGAGACTACACAGAACCCGGGCCAGACC
TACATGTGGTACACCGGCGAGCCCGTCTACGCCTTCGGCCACGGCCTGTTCTACACCACCTTCAACGCTT
CCTCAGCCCAAGCAGCGAAGACGAAGTATACCTTCAACATCACCGACCTCACCTCCGCCGCACACCCAGA
CACCACGACCGTCGGCCAACGCACCCTCTTCAACTTCACAGCCTCCATCACGAACTCCGGACAGAGGGAT
TCCGATTACACCGCCCTGGTGTACGCCAACACCTCGACTGCGGGCCCCTCCCCGTACCCGAATAAATGGC
TCGTCGGGTTCGACCGGCTCGCCGCCGTGGCGAAGGAGGGCGGCACGGCCGAGTTGAATGTGCCGGTGGC
GGTGGATCGGTTGGCGAGGGTGGATGAAGCGGGTAACACCGTGCTGTTTCCGGGGCGGTATGAGGTGGCC
CTGAATAATGAGCGCGAGGTCGTGGTCGAGGTGGAGTTGGTGGGTGAGCAGGTTGTGCTGTTGAAGTGGC
CGGAGGAGGTGCAGGGGGTGGCGGGGGATGAGTAG

FIGURE 35A

SEQ ID NO:40
Protein sequence of XlnD

mavaalalla llpqalgqhn ssyvdynvea npdlfpqcld tislsfpdcq sgplsknlvc
dstaspydra aalvslftle eliantgnts pgvprlglpp yqvwsealhg laranftdng
ayswatsfps pilsaaafnr tlinqiasii stqgrafnna grfgldvysp nintfrhpvw
grgqetpged aytltaayay eyitgiqggv npehlklaat akhfagydie nwdnhsrlgn
dvnitqqdla eyytpqflva ardahvhsfm csynavngvp scsntfflqt llrdtfsfvd
hgyvsgdcga vygvfnphgy aanepsaaad ailagtdidc gtsyqyhfne sittgavard
diergfirly anlvelgyfd gnssssnpyr slgwpdvqkt dawnisyeaa vegivllknd
gtlplaspse gknksialig pwanattqlq gnyygdapyl ispvdaftaa gytvhyapgt
eistnstanf saalsaaraa dtivflggid ntieaeaqdr ssiawpgnql elisqlaaqk
sddqplvvyq mgggqvdssa lksnakvnal lwggypgqsg glalrdiltg arapagrltt
tqypaayaes fsaldmnlrp nettqnpgqt ymwytgepvy afghglfytt fnassaqaak
tkytfnitdl tsaahpdttt vgqrtlfnft asitnsgqrd sdytalvyan tstagpspyp
nkwlvgfdrl aavakeggta elnvpvavdr larvdeagnt vlfpgryeva lnnerevvve
velvgeqvvl lkwpeevqgv agde

FIGURE 35B

SEQ ID NO:41
Nucleotide sequence of Fv30A, a GH30 enzyme from *Fusarium verticillioides*

atgctcttctcgctcgttcttcctacccttgcctttcaagccagcctggcgctcggcgatacatccgt
tactgtcgacaccagccagaaactccaggtcatcgatggctttggtgtctcagaagcctacggccacg
ccaaacaattccaaaacctcggtcctggaccacagaaagagggcctcgatcttctcttcaacactaca
accggcgcaggcttatccatcatccgaaacaagatcggctgcgacgcctccaactccatcaccagcac
caacaccgacaacccagataagcaggctgtttaccattttgacggcgatgatgatggtcaggtatggt
ttagcaaacaggccatgagctatggtgtagatactatctacgctaatgcttggtctgcgcctgtatac
atgaagtcagcccagagtatgggccgtctctgcggtacacctggtgtgtcgtgctcctctggagattg
gagacatcgttacgttgagatgatagctgagtacctctcctactacaagcaggctggcatcccagtgt
cgcacgttggattcctcaatgagggtgacggctcggactttatgctctcaactgccgaacaggctgca
gatgtcattcctcttctacacagcgctttgcagtccaagggccttggcgatatcaagatgacgtgctg
tgataacatcggttggaagtcacagatggactataccgccaagctggctgagcttgaggtggagaagt
atctatctgtcatcacatcccacgagtactccagcagccccaaccagcctatgaacactacattgcca
acctggatgtccgagggagctgccaatgaccaggcatttgccacagcgtggtacgtcaacggcggttc
caacgaaggtttcacatgggcagtcaagatcgcacaaggcatcgtcaatgccgacctctcagcgtata
tctactgggagggcgttgagaccaacaacaaggggtctctatctcacgtcatcgacacggacggtacc
aagtttaccatatcctcgattctctgggccattgctcactggtcgcgccatattcgccctggtgcgca
tagactttcgacttcaggtgttgtgcaagatacgattgttggtgcgtttgagaacgttgatggcagtg
tcgtcatggtgctcaccaactctggcactgctgctcagactgtggacctgggtgtttcgggaagtagc
ttctcaacagctcaggctttcacttcggatgctgaggcgcagatggtcgataccaaggtgactctgtc
cgacggtcgtgtcaaggttacggtcccggtgcacggtgtcgtcactgtgaagctcacaacagcaaaaa
gctccaaaccggtctcaactgctgtttctgcgcaatctgcccccactccaactagtgttaagcacacc
ttgactcaccagaagacttcttcaacaacactctcgaccgccaaggccccaacctccactcagactac
ctctgtagttgagtcagccaaggcggtgaaataccctgtcccccctgtagcatccaagggatcctcga
agagtgctcccaagaagggtaccaagaagaccactacgaagaagggctcccaccaatcgcacaaggcg
catagtgctactcatcgtcgatgccgccatggaagttaccgtcgtggccactgcaccaactaa

FIGURE 36A

SEQ ID NO:42
Protein sequence of Fv30A

mlfslvlptlafqaslalgdtsvtvdtsqklqvidgfgvseayghakqfqnlgpgpqkegldllfntt
tgaglsiirnkigcdasnsitstntdnpdkqavyhfdgdddgqsaqsmgrlcgtpgvscssgdwrhry
vemiaeylsyykqagipvshvgflnegdgsdfmlstaeqaadvipllhsalqskglgdikmtccdnig
wksqmdytaklaelevekylsvitsheyssspnqpmnttlptwmsegaandqafatawyvnggsnegf
twavkiaqgivnadlsayiywegvetnnkgslshvidtdgtkftissilwaiahwsrhirpgahrlst
sgvvqdtivgafenvdgsvvmvltnsgtaaqtvdlgvsgssfstaqaftsdaeaqmvdtkvtlsdgrv
kvtvpvhgvvtvklttaksskpvstavsaqsaptptsvkhtlthqktssttlstakaptstqttsvve
sakavkypvppvaskgssksapkkgtkktttkkgshqshkahsathrrcrhgsyrrghctn

FIGURE 36B

SEQ ID NO:43
Nucleotide sequence of Fv30B, a GH30 enzyme from *Fusarium verticillioides*

atgaatcctttatctctcggccttgccgccttgagccttctgggctacgtgggtgtcaactttgttgc
agccttccccacggattcaaactcaggctccgaagtcttgatttctgtcaatggccacgttaaacacc
aagagcttgacggatttggtgcttcacaagcattccaacgggccgaagacattcttggaaaagacggt
ctgtccaaagaagggactcagcatgtactggacttgctgttcagcaaggatatcggtgcgggcttctc
tatcctgcgtaatggcattggctcaagcaacagttctgacaagaacttcatgaattcaatcgagccat
tctcgccaggctcacccggagcaaagccacactacgtctgggatggctatgatagcggacaactcacc
gtcgctcaagaagcattcaagagaggattgaagttcctctatggcgatgcttggtccgctcctggtta
catgaagacaaaccacgatgagaataacgggggggtatttgtgtggtgttacaggtgctgcctgcgctt
ctggcgactggaagcaggcttacgcagactacttgctgcagtgggttgagttctaccgcaagtcaggc
gtcaaggtcaccaacctaggattccttaacgagcctcagttcgccgctccctacgccggcatgctgtc
taacggcacacaggctgccgacttcatacgtgtactgggcaagacaatcagaaaacgaggtatccacg
accttacaatcgcctgctgtgatggcgagggctgggatctccaagaagatatgatggctggtttgact
gctggacctgatccggcaatcaactacctcagtgtcgttactgggcacggctacgtttcaccaccgaa
ccatccgctttcaacaacaaagaagacgtggctcaccgagtgggctgatctcacaggccagttcacgc
cctacacgttctacaacaatagcggtcaggggggaaggtatgacctgggctggccgtatccagacggcg
cttgtagatgccaatgtcagcggctttctctattggatcggagccgagaactcgaccaccaacagtgc
tctgatcaacatgatcggcgacaaggtcatcccttccaagaggttctgggcctttgcatccttcagtc
ggtttgctagacctggtgctcgtcgcattgaagccacgagctccgttcctctggtcacagtcagttca
tttctgaataccgacggtactgtcgcgacgcaggtgctgaacaacgacacggttgctcacagtgtgca
actcgttgtctctggcacaggtcgaaatcctcatagcttgaagccgtttttgaccgataattctaatg
atttgactgccttgaagcatttgaaggctactggaaagggttcatttcagactacgattcctcctcga
tctcttgttagctttgttacagatttctaacaaagacaatattacttgaagaagacgactatgagggc
tgctttgatcaagttgactatgtctagtatgttggtgtaaatctcctaacaatcttgttgggctgctt
atttcggcttagttacgcaacgtcatgttcagtgtgccgaaagccgaaccacgaaaatagctcacaag
accattctggattttgacacgataagatcctgccttttttttcatacttgttcctctctttcacttgg
cgaaatatgctgttttacgtatccatgctctccc

FIGURE 37A

SEQ ID NO:44
Protein sequence of Fv30B
mnplslglaalsllgyvgvnfvaafptdsnsgsevlisvnghvkhqeldgfgasqafqraedilgkdg
lskegtqhvldllfskdigagfsilrngigssnssdknfmnsiepfspgspgakphyvwdgydsgqlt
vaqeafkrglkflygdawsapgymktnhdennggylcgvtgaacasgdwkqayadyllqwvefyrksg
vkvtnlgflnepqfaapyagmlsngtqaadfirvlgktirkrgihdltiaccdgegwdlqedmmaglt
agpdpainylsvvtghgyvsppnhplsttkktwltewadltgqftpytfynnsgqgegmtwagriqta
lvdanvsgflywigaensttnsalinmigdkvipskrfwafasfsrfarpgarrieatssvplvtvss
flntdgtvatqvlnndtvahsvqlvvsgtgrnphslkpfltdnsndltalkhlkatgkgsfqttippr
slvsfvtdf*

FIGURE 37B

SEQ ID NO:45
Nucleotide sequence of Fv39A, a GH39 enzyme from *Fusarium verticillioides*

atgcactacgctaccctcaccactttggtgctggctctgaccaccaacgtcgctgcacagcaaggcac
agcaactgtcgacctctccaaaaatcatggaccggcgaaggcccttggttcaggcttcatatacggct
ggcctgacaacggaacaagcgtcgacacctccataccagatttcttggtaactgacatcaaattcaac
tcaaaccgcggcggtggcgcccaaatcccatcactgggttgggccagaggtggctatgaaggatacct
cggccgcttcaactcaaccttatccaactatcgcaccacgcgcaagtataacgctgactttatcttgt
tgcctcatgacctctggggtgcggatggcgggcagggttcaaactccccgtttcctggcgacaatggc
aattggactgagatggagttattctggaatcagcttgtgtctgacttgaaggctcataatatgctgga
aggtcttgtgattgatgtttggaatgagcctgatattgatatcttttgggatcgcccgtggtcgcagt
ttcttgagtattacaatcgcgcgaccaaactacttcggtgagtctactactgatccatacgtatttac
agtgagctgactggtcgaattagaaaaacacttcccaaaactcttctcagtggcccagccatggcaca
ttctcccattctgtccgatgataaatggcatacctggcttcaatcagtagcgggtaacaagacagtcc
ctgatatttactcctggcatcagattggcgcttgggaacgtgagccggacagcactatccccgacttt
accaccttgcgggcgcaatatggcgttcccgagaagccaattgacgtcaatgagtacgctgcacgcga
tgagcaaaatccagccaactccgtctactacctctctcaactagagcgtcataaccttagaggtcttc
gcgcaaactggggtagcggatctgacctccacaactggatgggcaacttgatttacagcactaccggt
acctcggaggggacttactaccctaatggtgaatggcaggcttacaagtactatgcggccatggcagg
gcagagacttgtgaccaaagcatcgtcggacttgaagtttgatgtctttgccactaagcaaggccgta
agattaagattatagccggcacgaggaccgttcaagcaaagtataacatcaaaatcagcggtttggaa
gtagcaggacttcctaagatgggtacggtaaaggtccggacttatcggttcgactgggctgggccgaa
tggaaaggttgacgggcctgttgatttgggggagaagaagtatacttattcggccaatacggtgagca
gcccctctacttga

FIGURE 38A

SEQ ID NO:46
Protein sequence of Fv39A

mhyatlttlvlalttnvaaqqgtatvdlsknhgpakalgsgfiygwpdngtsvdtsipdflvtdikfn
snrgggaqipslgwarggyegylgrfnstlsnyrttrkynadfillphdlwgadggqgsnspfpgdng
nwtemelf**wnqlvsdlkahnmleglvidvwnepdidifwdrpwsqfleyynratkllrktlpktllsg**
pamahspilsddkwhtwlqsvagnktvpdiyswhqigawerepdstipdfttlraqygvpekpidvne
yaardeqnpansvyylsqlerhnlrglranwgsgsdlhnwmgnliysttgtsegtyypngewqaykyy
aamagqrlvtkassdlkfdvfatkqgrkikiiagtrtvqakynikisglevaglpkmgtvkvrtyrfd
wagpngkvdgpvdlgekkytysantvsspst

FIGURE 38B

SEQ ID NO:47
Nucleotide sequence of Fv39B, a GH39 enzyme from *Fusarium verticillioides*

```
atgtggttaatcaaggcctgttccgtcctcgccgctctctccactgtagctgctgacagccccggtcc
caccatcgacttctcctccaacactggagagcctcagcatctcgctgctggtatcctgtacggtatac
ccgacgatgggaaccagatcccagatgatcttctctctggctttggcttcaactactatcgcggtgca
ggtgcccaagtctctcatggatggagttatgacgaggctggcttccagcagcgttttgaaagcgcgca
taacaactacatcgtcacgcgtcgtcacaacggcggctttgtcttgttgctgaatgacctctggggct
ttgattgttcttctaacaacgatacctcacctggtccaggcgataatggcgattggtcgtcctatgac
aagttcgttcaggcgattattgccaatgtcaagaagtacaacatgcaggaaggcttggtcattgatat
ctggaacgagccagaggggggctgtttctggggccgtagcattgaccaatggcttcagatgtggggtc
gcggctggcatcaattcaagtaagtactagatacctctgaggacggatgggacaagaactgactgtct
attcagtgatgccttcggggacagcgtgttgacatccggaccaactcttgcaggcgagccgggaacaa
acgatgactggtggacccaatgggcccaattcgtcaagaacaacgactccatccccgaccaatacgca
tggcacgaggaaggaggctcaggttccaacttcgagaacagctacggcgtcctgcaacaaattctcac
taaatacggtcttccccaacgccaaatcaacatcaacgaatacgctacgttcaatgaacaagtccccg
ccggttctgccttctggatctcccagttcgagcgccgtaatgctatcggtcttcgcggcaattggcta
ggaggcactcaacttcacgatctggccgctagtcttctgtccaagcctgacccctcggactacgcttc
cacgggatactttgccaatggagactggtgggtgtataactactactctcacaacatgacgggacagc
gcgtttcgacttcggtgtcttccgatggaaggctggatgcttatgcaacggtggatactacggcgcgc
acggctagagtattgcttggctgccatccgcctacgactggtacttatgatgtgacattctctggtct
gacaaagttgggtctgccatcttctggaacacttcaggttaggacttggaagtttgctgtgggcagtg
atgtgcattacagccaggtgggttctcctcaggatctgggtaactatggtcacactattagcaacggt
caggttaccttgccgttctatcagactgatgatgtgactacatacgcatgggagttcaaattctag
```

FIGURE 39A

SEQ ID NO:48
Protein sequence of Fv39B

```
mwlikacsvlaalstvaadspgptidfssntgepqhlaagilygipddgnqipddllsgfgfnyyrga
gaqvshgwsydeagfqqrfesahnnyivtrrhnggfvlllndlwgfdcssnndtspgpgdngdwssyd
kfvqaiianvkkynmqeglvidiwnepeggcfwgrsidqwlqmwgrgwhqfndafgdsvltsgptlag
epgtnddwwtqwaqfvknndsipdqyawheeggsgsnfensygvlqqiltkyglpqrqinineyatfn
eqvpagsafwisqferrnaiglrgnwlggtqlhdlaasllskpdpsdyastgyfangdwwvynyyshn
mtgqrvstsvssdgrldayatvdttartarvllgchppttgtydvtfsgltklglpssgtlqvrtwkf
avgsdvhysqvgspqdlgnyghtisngqvtlpfyqtddvttyawefkf*
```

FIGURE 39B

SEQ ID NO:49
Nucleotide sequence of XynB, a GH39 enzyme from *Thermoanaerobacter saccharolyticum*

```
atggtaaaaa taaagatacc aaaaaattct gatggcaaaa aattcaccag tagatggaga
tattgtgtag gtacaggaag gttgggactt gcgctgcaaa aagagtacat ggatacttta
aaatttgtga aagaaaatat agacttcaag tatataagag gacatggcct tttgtgcgac
gatgtaggta tttaccgtga ggacgtagta ggcaacgatg taaggccatt ttacaatttc
acgtatatag atagaatctt tgattcattt ttggaattag ggataaggcc atttgttgaa
gtcggattta tgcctaaaag attagcatct ggtacacaga cagtatttta ttgggaggga
aatgtcaccc ctcccaaaga ttatgaaaaa tggagcaacc ttataaaagc tgttgtttca
cattttatat caaggtatgg catagatgaa gtcgtaaaat ggccatttga aatatggaat
gagccaaacc taaaagagtt ttggaaagat gctgatgaga aagaatactt caagctgtac
aaggttactg caaaggcgat taaagaagta aatgagaatt tgcaggtagg aggacctgct
atatgtggtg gtgctgacta ctggatagaa gatttttga atttctgcta tgaagaaaat
gttcctgttg attttgtatc gcgacacgca tatacgtcta agcaaggtga atatacgcca
catctcatat accaggagat tatgccatct gaatacatgc taaacgaatt caaaacagtg
agagagatca taaaaaactc acattttccg aaccttccgt ttcatataac ggagtacaat
acatcttaca gtccattaaa tcctgtacat gatacgcctt ttaatgcggc gtatcttgcg
aggattttaa gtgaaggcgg agattatgtt gattcatttt cctattggac gtttagcgat
gtgtttgaag aaagagatgt gccaagatcg cagtttcatg gaggatttgg actagttgca
ttgaataaga taccaaagcc gacttttcac atgtttaaat ttttcaatgc tatgggagaa
gaggtgcttt acagagataa ccatatgctt ataactagaa gggatgatgg gtcgattgca
ttgattgctt ggaatgagat aatggagaaa acagaaaatc cagataagga atatgaactg
gaaatacctg taggattcaa agatgtcttt ataaagaaac agatgataga tgaggatcac
ggcaatcctt ggggtacgtg gatacacatg ggaaggccga gattcccaag taaagaacaa
attaagactt taagggatat tgcaaagcct aaaatcaaaa caggtagagc cacatcaaat
gatggctatg taaatttgaa atttagattg gggaaaaatg ctgtggtatt gtttgaattg
actgaagtaa tggatgaatc aaacacttat ataggacttg atgatagcaa gataaacgga
tattga
```

FIGURE 40A

SEQ ID NO:50
Protein sequence of XynB

```
MIKVRVPDFS DKKFSDRWRY CVGTGRLGLA LQKEYIETLK YVKENIDFKY IRGHGLLCDD
VGIYREDVVG DEVKPFYNFT YIDRIFDSFL EIGIRPFVEI GFMPKKLASG TQTVFYWEGN
VTPPKDYEKW SDLVKAVLHH FISRYGIEEV LKWPFEIWNE PNLKEFWKDA DEKEYFKLYK
VTAKAIKEVN ENLKVGGPAI CGGADYWIED FLNFCYEENV PVDFVSRHAT TSKQGEYTPH
LIYQEIMPSE YMLNEFKTVR EIIKNSHFPN LPFHITEYNT SYSPQNPVHD TPFNAAYIAR
ILSEGGDYVD SFSYWTFSDV FEERDVPRSQ FHGGFGLVAL NMIPKPTFYT FKFFNAMGEE
MLYRDEHMLV TRRDDGSVAL IAWNEVMDKT ENPDEDYEVE IPVRFRDVFI KRQLIDEEHG
NPWGTWIHMG RPRYPSKEQV NTLREVAKPE IMTSQPVAND GYLNLKFKLG KNAVVLYELT
ERIDESSTYI GLDDSKINGY
```

FIGURE 40B

SEQ ID NO:51

Nucleotide sequence of XylA, A GH52 enzyme from *Geobacillus stearothermophilus*

```
atgccaacca atgtattttt caacgcccat cactcgccgg ttggggcgtt tgccagcttt
acgctagggt ttccgggaaa aagcggagga ctggacttgg aactcgcccg accgccacgg
caaaatgtct ttattggcgt tgagtcgccg catgagccgg ggctgtatca tatccttcca
ttcgcggaaa cagcaggcga ggatgaaagc aaacgatatg acattgaaaa tcctgatccg
aatccgcaaa aaccaaacat tctcatcccg tttgcgaaag aggagatcaa gcgcgaattc
tgtgtggcaa cggatacatg gaaagctggg gatttaacgt ttacgattta ttcgccggta
aaggcggtgc ctgatcccga aacagcggcc gaggaagaac tcaagttggc gttggtccca
gctgtcattg tcgagatgac gatcgataac acgaacggaa caagaacacg acgggcgttt
ttcggattcg aaggcaccga cccgtatacc tcgatgcggc ggatcgatga tacatgcccg
cagctgcgcg gggtcggtca agggcggatt ttgggcattg tatccaagga tgagggtgtt
cgctcagcgc tgcattttag catggaggat atcttaacgg cgactctcga agaaaactgg
acgtttgggc ttgggaaagt cggtgcgtta atcgttgatg tgccggcggg agaaaagaaa
acgtatcaat ttgctgtttg tttttaccgc ggtggttatg ttaccgcagg aatggatgcc
tcttattttt acacccgttt cttccataat atcgaagaag tcggtcttta tgcgttagag
caggccgagg tgttaaagga gcaggcgttc cgttcgaatg aactcattga aaaagaatgg
ctctccgatg atcaaaagtt tatgatggcg cacgcgatcc gcagctacta tggcaataca
caactgcttg agcatgaagg aaagccgatt tgggtcgtta atgaaggcga gtaccggatg
atgaatacgt ttgatctcac cgtcgaccag ctctttttg agttgaaaat gaatccgtgg
acggtgaaaa atgtgcttga cttctatgtc gagcgctaca gctatgagga tcgtgtccgt
ttcccaggag atgagacgga atatcccggc ggcatcagct tcactcatga tatgggagtc
gctaacacgt tctcgcgtcc gcattactcg tcatatgagc tatacggaat cagcggctgc
ttttcgcata tgacgcacga acagctcgtc aactgggtgc tttgcgcggc ggtatacatc
gaacaaacga aagactgggc atggcgcgac cggcggctta cgatcttgga acaatgtctc
gaaagcatgg tgcgtcgtga tcatccggat ccagaaaagc ggaacggcgt gatggggctt
gacagcaccc gcaccatggg tggagcggaa atcacaacgt atgatagttt ggatgtttcc
ctcggccagg cgcgcaacaa tttatatttg gcaggaaaat gttgggctgc ctatgtggcg
ctcgaaaagt tgttccgcga tgtcggcaaa gaagaactgg ctgcattggc aagggagcag
gcggaaaaat gcgccgcgac gattgtcagt cacgtgacgg aggacgggta tatcccagcc
gtgatgggag aaggaaatga ctcgaaaatc attccggcta ttgaggggct tgtgtttcct
tactttacga actgccatga ggcgttaaga gaagacggac gttttggaga ctatattcgt
gcactgcgac aacatttgca atatgtgttg cgggaaggaa tttgcctatt cccggacggg
ggatggaaaa tttcctcgac aagcaacaac tcgtggttga gcaaaattta cttatgccag
tttattgccc gccgcatttt agggtgggaa tgggatgaac aaggaaaacg agctgatgcg
gctcatgttg cgtggctcac gcatccgacg ctctccattt ggagttggag cgaccaaatt
atcgctggcg aaatcagcgg cagcaaatac tacccgcgcg gcgtgacgag cattttatgg
ttggaggagg gggaatga
```

FIGURE 41A

SEQ ID NO:52
Protein sequence of XylA

mptnlffnah hspvgafasf tlgfpgksgg ldlelarppr qnvligvesl hesglyhvlp
fletaeedes krydienpdp npqkpnilip fakeeiqref hvatdtwkag dltftiyspv
kavpnpetad eeelklalvp avivemtidn tngtrarraf fgfegtdpyt smrriddtcp
qlrgvgqqgri lsivskdegv rsalhfsmed iltaqleenw tfglgkvgal ivdvpagekk
tyqfavcfyr ggyvtagmda syfytrffqn ieevglyale qaevlkeqsf rsnkliekew
lsddqtfmma hairsyygnt qllehegkpi wvvnegeyrm mntfdltvdq lffelklnpw
tvknvldlyv erysyedrvr fpgeeteyps gisfthdmgv antfsrphys syelygisgc
fshmtheqlv nwvlcaavyi eqtkdwawrd krlaileqcl esmvrrdhpd peqrngvmgl
dstrtmggae ittydsldvs lgqarnnlyl agkcwaayva leklfrdvgk eelaalageq
aekcaativs hvtddgyipa imgegndski ipaieglvfp yftnchealD engrfgayiq
alrnhlqyvl regiclfpdg gwkisstsnn swlskiylcq fiarhilgwe wdeqgkrada
ahvawlthpt lsiwswsdqi iageitgsky yprgvtsilw leege

FIGURE 41B

SEQ ID NO:53

Nucleotide sequence of Xyl1, a GH54 enzyme from *Trichoderma koningii (Hypocrea koningii)*

```
ATGCTCTCCAACGCTCGTATCATCGCAGCGGGCTGTATTGCTGCAGGCTCTCTCGTTGCT
GCTGGGCCTTGTGACATCTACTCCTCCGGCGGCAACGCCTTGCGTTGCCGCCCACAGCACC
ACTCGAGCTCTGTTCAGCGCTTATACCGGCCCGTTATACCAGGTAAAGCGCGGCTCCGAT
GGTGCCACAACCGCCATATCGCCCCTCTCAAGTGGTGTGGCCAACGCTGCCGCTCAAGAT
GCTTTCTGTGCGGGAACTACATGCCTCATTACCATCATATACGACCAGTCGGGTCGCGGC
AACCATCTCAGGGAGGCCCCGCCGGGCGGCTTCAGCGGCCCGGAATCCAACGGCTATGAC
AACCTGGCTAGTGCAATTGGGGCGCCCGTAACACTCAACGGCCAGAAGGCGTATGGAGTT
TTCGTGTCTCCAGGAACGGGGTATCGGAATAACGCTGCCAGCGGCACAGCCAAAGGAGAT
GCCGCGGAGGGCATGTATGCGGTTCTCGATGGTACACACTACAACGGCGCCTGCTGCTTT
GACTATGGCAACGCCGAGACCAACAGCCGCGATACAGGCAACGGTCATATGGAGGCCATC
TATTTTGGCGACAGCACTGTCTGGGGTACTGGCTCAGGCAAGGGTCCGTGGATCATGGCT
GATCTCGAGAACGGCTTGTTCTCAGGCTCCAGTCCCGGCAACAATGCCGGTGATCCGTCC
ATCTCGTACCGGTTCGTCACTGCAGCGATCAAGGGGCAGCCAAACCAATGGGCAATCCGT
GGCGGCAATGCTGCGTCTGGCTCGCTGTCAACTTTCTACAGCGGCGCTCGCCCACAAGTC
TCCGGATACAATCCGATGAGCAAAGAGGGCGCCATCATTCTCGGCATTGGCGGCGACAAC
AGCAACGGCGGCCAGGGCACATTCTATGAGGGCGTCATGACCTCTGGATATCCTTCCGAT
GCAACAGAGAATTCAGTGCAAGCCAACATCGTAGCTGCCA
```

FIGURE 42A

SEQ ID NO:54

Protein sequence of Xyl1

```
MLSNARIIAA GCIAAGSLVA AGPCDIYSSG GTPCVAAHST TRALFSAYTG PLYQVKRGSD
GATTAISPLS SGVANAAAQD AFCAGTTCLI TIIYDQSGRG NHLREAPPGG FSGPESNGYD
NLASAIGAPV TLNGQKAYGV FVSPGTGYRN NAASGTAKGD AAEGMYAVLD GTHYNGACCF
DYGNAETNSR DTGNGHMEAI YFGDSTVWGT GSGKGPWIMA DLENGLFSGS SPGNNAGDPS
ISYRFVTAAI KGQPNQWAIR GGNAASGSLS TFYSGARPQV SGYNPMSKEG AIILGIGGDN
SNGGQGTFYE GVMTSGYPSD ATENSVQANI VAARYAVAPL TSGPALTVGS SISLRATTAC
CTTRYIAHSG STVNTQVVSS SSATALKQQA SWTVRAGLAN NACFSFESQD TSGSYIRHSN
FGLVLNANDG SKLFAEDATF CTQAGINGQG SSIRSWSYPT RYFRHYNNTL YIASNGGVHV
FDATAAFNDD VSFVVSGGFA
```

FIGURE 42B

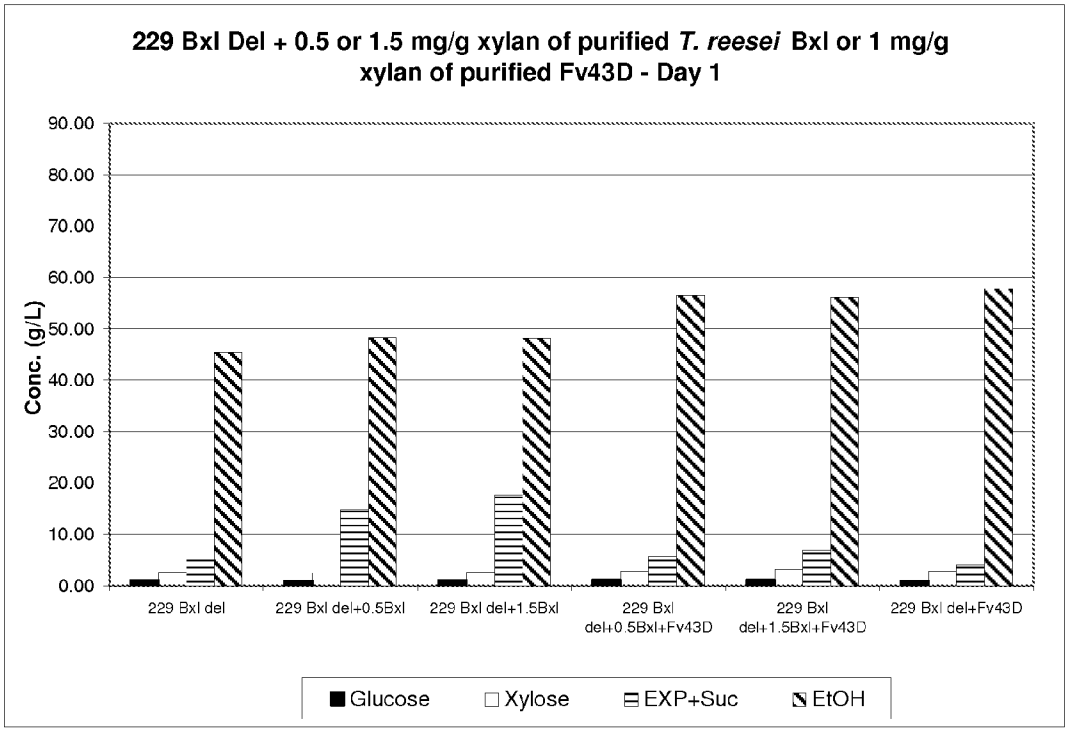

229 Bxl Del + 0.5 or 1.5 mg/g xylan of purified T. reesei Bxl or 1 mg/g xylan of purified Fv43D - Day 1
Conc. (g/L)
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
229 Bxl del
229 Bxl del+0.5Bxl
229 Bxl del+1.5Bxl
229 Bxl del+0.5Bxl+Fv43D
229 Bxl del+1.5Bxl+Fv43D
229 Bxl del+Fv43D
Glucose
Xylose
EXP+Suc
EtOH

FIGURE 43A

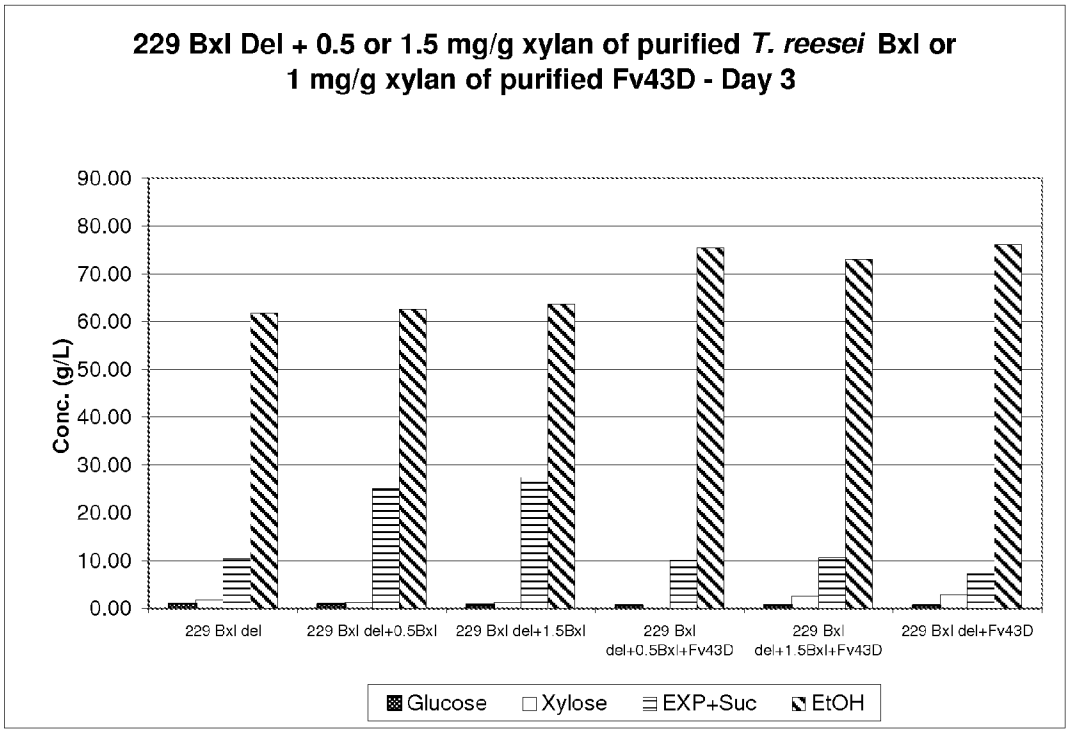

229 BxI Del + 0.5 or 1.5 mg/g xylan of purified T. reesei BxI or 1 mg/g xylan of purified Fv43D - Day 3
Conc. (g/L)
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
229 BxI del
229 BxI del+0.5BxI
229 BxI del+1.5BxI
229 BxI del+0.5BxI+Fv43D
229 BxI del+1.5BxI+Fv43D
229 BxI del+Fv43D
Glucose
Xylose
EXP+Suc
EtOH

FIGURE 43B