

(19) **DANMARK**



Patent- og
Varemærkestyrelsen

(10) **DK/EP 2480660 T5**

(12) **Rettet oversættelse af
europæisk patentskrift**

-
- (51) Int.Cl.: **C 12 N 9/24 (2006.01)** **C 12 P 7/04 (2006.01)** **C 12 P 7/08 (2006.01)**
C 12 P 19/00 (2006.01)
- (45) Oversættelsen bekendtgjort den: **2020-11-09**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2020-08-05**
- (86) Europæisk ansøgning nr.: **10777132.1**
- (86) Europæisk indleveringsdag: **2010-09-22**
- (87) Den europæiske ansøgnings publiceringsdag: **2012-08-01**
- (86) International ansøgning nr.: **US2010049849**
- (87) Internationalt publikationsnr.: **WO2011038019**
- (30) Prioritet: **2009-09-23 US 245273 P** **2009-12-23 US 289886 P**
- (84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO SE SI SK SM TR**
- (73) Patenthaver: **Danisco US Inc., 925 Page Mill Road, Palo Alto, California 94304, USA**
- (72) Opfinder: **BOWER, Benjamin, Danisco US INC., 925 Page Mill Road, Palo Alto, California 94304, USA**
HSI, Megan Yee, Danisco US INC., 925 Page Mill Road, Palo Alto, California 94304, USA
KAPER, Thijs, Danisco Us Inc., 925 Page Mill Road, Palo Alto, California 94304, USA
KELEMEN, Bradley R., Danisco US INC., 925 Page Mill Road, Palo Alto, California 94304, USA
LANTZ, Suzanne E., Danisco US INC., 925 Page Mill Road, Palo Alto, California 94304, USA
LARENAS, Edmund A., Danisco US INC., 925 Page Mill Road, Palo Alto, California 94304, USA
MITCHINSON, Colin, Danisco US Inc., 925 Page Mill Road, Palo Alto, California 94304, USA
KIM, Steven, Danisco US INC., 925 Page Mill Road, Palo Alto, California 94304, USA
HITZ, William D., 404 Hillside Road, Wilmington, Delaware 19807, USA
EMPTAGE, Mark, 2822 Landon Drive, Wilmington, Delaware 19810, USA
WING, Keith Dumont, 122 Yardley Lane, Wilmington, Delaware 19810, USA
- (74) Fuldmægtig i Danmark: **Marks & Clerk (Luxembourg) LLP, 44 rue de la Vallée, B.P. 1775, L-1017 Luxembourg, Luxembourg**
- (54) Benævnelse: **HIDTIL UKENDTE GLYCOSYLHYDROLASEENZYMER OG ANVENDELSER HERAF**
- (56) Fremdragne publikationer:
WO-A2-2008/045977
WO-A2-2009/033071
WO-A2-2009/045627
US-A1- 2009 050 134
SORENSEN H R ET AL: "Enzymatic hydrolysis of water-soluble wheat arabinoxylan. 1. Synergy between alpha-L-arabinofuranosidases, endo-1,4-beta-xylanases, and beta-xylosidase activities", BIOTECHNOLOGY AND BIOENGINEERING, WILEY & SONS, HOBOKEN, NJ, US, vol. 81, no. 6, 20 March 2003 (2003-03-20) , pages 726-

Fortsættes ...

731, XP002386252, ISSN: 0006-3592, DOI: DOI:10.1002/BIT.10519

SAHA B C: "Xylanase from a newly isolated *Fusarium verticillioides* capable of utilizing corn fiber xylan.", *APPLIED MICROBIOLOGY AND BIOTECHNOLOGY* SEP 2001 LNKD- PUBMED:11601627, vol. 56, no. 5-6, September 2001 (2001-09) , pages 762-766, XP002633642, ISSN: 0175-7598

DATABASE EMBL [Online] 12 July 2005 (2005-07-12), "EST1021346 *FvI* *Gibberella moniliformis* cDNA clone FVIE185, mRNA sequence.", XP002633643, retrieved from EBI accession no. EMBL:DR631218 Database accession no. DR631218 -& BROWN D W ET AL: "Comparative analysis of 87,000 expressed sequence tags from the fumonisin-producing fungus *Fusarium verticillioides*", *FUNGAL GENETICS AND BIOLOGY*, SAN DIEGO, CA, US, vol. 42, no. 10, 1 October 2005 (2005-10-01), pages 848-861, XP027195855, ISSN: 1087-1845 [retrieved on 2005-09-15]

DATABASE EMBL [Online] 15 April 2003 (2003-04-15), "Penicillium herquei mRNA for xylosidase, complete cds.", XP002633644, retrieved from EBI accession no. EMBL:AB093564 Database accession no. AB093564

DATABASE EMBL [Online] 19 April 2007 (2007-04-19), "Penicillium purpurogenum alpha-L-arabinofuranosidase 2 (abf2) gene, complete cds.", XP002633645, retrieved from EBI accession no. EMBL:EF490448 Database accession no. EF490448

MARGOLLES-CLARK E ET AL: "Expression patterns of ten hemicellulase genes of the filamentous fungus *Trichoderma reesei* on various carbon sources", *JOURNAL OF BIOTECHNOLOGY*, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 57, no. 1-3, 16 September 1997 (1997-09-16), pages 167-179, XP004097715, ISSN: 0168-1656, DOI: DOI:10.1016/S0168-1656(97)00097-7

MACH-AIGNER ASTRID R ET AL: "Transcriptional regulation of *xyr1*, encoding the main regulator of the xylanolytic and cellulolytic enzyme system in *Hypocrea jecorina*.", *APPLIED AND ENVIRONMENTAL MICROBIOLOGY* NOV 2008 LNKD- PUBMED:18791032, vol. 74, no. 21, November 2008 (2008-11), pages 6554-6562, XP002633646, ISSN: 1098-5336

DATABASE REFSEQ [Online] 9 April 2008 (2008-04-09), NCBI Reference Sequence (RefSeq) collection: "Hypothetical protein FG03609.1 [*Giberella zeae* PH-1]", XP002660306, Database accession no. XP_383785

DATABASE REFSEQ [Online] 9 April 2008 (2008-04-09), NCBI Reference Sequence (RefSeq) collection: "Hypothetical protein FG06463.1 [*Gibberella zeae* PH-1]", XP002660307, Database accession no. XP_386639.1

DESCRIPTION

2. TECHNICAL FIELD

[0001] The present disclosure generally pertains to glycosyl hydrolase enzymes, compositions comprising such enzymes, and methods of using the enzymes and compositions, for example for the saccharification or conversion of cellulosic and hemicellulosic materials into sugars.

3. BACKGROUND

[0002] Bioconversion of renewable lignocellulosic biomass to a fermentable sugar that is subsequently fermented to produce alcohol (e.g., ethanol) as an alternative to liquid fuels has attracted the intensive attention of researchers since the 1970s, when the oil crisis occurred because OPEC decreased the output of petroleum (Bungay, H. R., "Energy: the biomass options". NY: Wiley; 1981; Olsson L, Hahn-Hagerdal B. Enzyme Microb Technol 1996,18:312-31; Zaldivar, J et al., Appl Microbiol Biotechnol 2001, 56: 17-34; Galbe, M et al., Appl Microbiol Biotechnol 2002, 59:618-28). 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. Additionally, fermentable sugars are increasingly used to produce plastics, polymers and other biobased products, and this industry is expected to expand substantially in the coming years. Thus, the demand for abundant low cost fermentable sugars, which can be used as a feed stock in lieu of petroleum based feedstocks, continues to grow.

[0003] Chiefly among the useful renewable feedstocks are cellulose and hemicellulose (xylans), which can be converted into fermentable sugars. The enzymatic conversion of these polysaccharides to soluble sugars, for example glucose, xylose, arabinose, galactose, mannose, and/or other hexoses and pentoses, occurs due to combined actions of various enzymes. For example, endo-1,4- β -glucanases (EG) and exo-cellobiohydrolases (CBH) catalyze the hydrolysis of insoluble cellulose to celooligosaccharides (e.g., with cellobiose being a main product), while β -glucosidases (BGL) catalyzes the conversion of the oligosaccharides to glucose. Xylanases together with other accessory proteins (non-limiting examples of which include L- α -arabinofuranosidases, feruloyl and acetylxylan esterases, glucuronidases, and β -xylosidases) catalyze the hydrolysis of hemicelluloses.

[0004] The cell walls of plants are composed of a heterogeneous mixture of complex polysaccharides that interact through covalent and noncovalent means. Complex polysaccharides of higher plant cell walls include, for example, cellulose (β -1,4 glucan), which generally constitutes 35-50% of carbon found in cell wall components. Cellulose polymers self associate through hydrogen bonding, van der Waals interactions and hydrophobic interactions to form semi-crystalline cellulose microfibrils. These microfibrils also include noncrystalline

regions, generally known as amorphous cellulose. The cellulose microfibrils are embedded in a matrix formed of hemicelluloses (including, *e.g.*, xylans, arabinans, and mannans), pectins (*e.g.*, galacturonans and galactans), and various other β -1,3 and β -1,4 glucans. These matrix polymers are often substituted with, for example, arabinose, galactose and/or xylose residues to yield highly complex arabinoxylans, arabinogalactans, galactomannans, and xyloglucans. The hemicellulose matrix is, in turn, surrounded by polyphenolic lignin.

[0005] The complexity of the matrix makes it difficult to degrade by microorganisms as lignin and hemicellulose components must be broken down before enzymes can act on the core cellulose microfibrils. Ordinarily, a consortium of different enzymatic activities is required to break down cell wall polymers to release the constituent monosaccharides. For saccharification of plant cell walls, the lignin must be permeabilized and hemicellulose disrupted to allow cellulose-degrading enzymes to act on their substrate.

[0006] Regardless of the type of cellulosic feedstock, the cost and hydrolytic efficiency of enzymes are major factors that restrict the commercialization of biomass bioconversion processes. The production costs of microbially produced enzymes are tightly connected with the productivity of the enzyme-producing strain and the final activity yield in the fermentation broth. The hydrolytic efficiency of a multienzyme complex depends both on properties of individual enzymes, the synergies among them, and their ratio in the multienzyme blend.

[0007] US2006/050134 discloses a method for treating biomass was developed that uses an apparatus which moves a biomass and dilute aqueous ammonia mixture through reaction chambers without compaction.

[0008] WO2008/045977 discloses an ss-D-xylosidase with a k_{cat} value for catalysis of approximately 185 sec^{-1} for 1, 4-ss-D-xylobiose when measured at a pH of 5.3 and a temperature of 25°C . Also discussed is the use of the ss-D-xylosidase alone or in combination with other hydrolytic or xylanolytic enzymes for treatment of lignocellulosic or hemicellulosic plant materials or plant material hydrolysates or xylooligosaccharides.

[0009] SORENSEN H R et al., BIOTECHNOLOGY AND BIOENGINEERING, vol. 81, no. 6, 20 March 2003, pages 726-731, discusses synergy between α -L-arabinofuranosidases, endo-1,4- β -xylanases, and β -xylosidase activities of cellulytic enzyme preparations in the enzymatic hydrolysis of water-soluble wheat arabinoxylan.

[0010] SAHA BC, APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, vol. 56, no. 5-6, September 2001, pages 762-766, discloses a *Fusarium verticillioides* (NRRL 26518) fungus isolated by screening soil samples using corn fiber xylan as carbon source. The extracellular xylanase from this fungal strain was purified to apparent homogeneity from the culture supernatant by ultrafiltration using a 30,000 cut-off membrane, octyl-Sepharose chromatography and Bio gel A-0.5 m gel filtration. The purified xylanase (specific activity 492 U/mg protein; MW 24,000; pI 8.6) displayed an optimum temperature at 50°C and optimum pH at 5.5, a pH stability range from 4.0 to 9.5 and thermal stability up to 50°C . It hydrolyzed a

variety of xylan substrates mainly to xylobiose and higher short-chain xylooligosaccharides. No xylose was formed. The enzyme did not require metal ions for activity and stability.

[0011] There exists a need in the art to identify enzyme and/or enzymatic blends/compositions that are capable of converting plant and/or other cellulosic or hemicellulosic materials into fermentable sugars with improved efficacy and yield.

4. SUMMARY

[0012] The disclosure provides certain glycosyl hydrolase polypeptides having hemicellulolytic activity, including, *e.g.*, xylanases (*e.g.*, endoxylanases), xylosidases (*e.g.*, β -xylosidases), arabinofuranosidases (*e.g.*, L- α -arabinofuranosidases), nucleic acids encoding these polypeptides, and methods for making and using the polypeptides and/or nucleic acids. The disclosure is based, in part, on the discovery of novel enzymes and variants having xylanase, β -xylosidase, and/or L- α -arabinofuranosidase activities. The disclosure is also based on the identification of enzyme blends (or compositions) that efficiently catalyze the hydrolysis of cellulosic and hemicellulosic materials. For purpose of this disclosure, an enzyme can be defined either as a polypeptide having the particular enzymatic activity or as that enzyme. For example, a xylanase can be referred to as a polypeptide having xylanase activity or as a xylanase enzyme, and a β -xylosidase can be referred to as either a polypeptide having β -xylosidase activity or a β -xylosidase enzyme.

[0013] The enzymes and/or enzyme blends/compositions of the disclosure can be used to produce sugars from biomass. The sugars so produced can be used by microorganisms for ethanol production or can be used to produce other bioproducts in various industrial applications. Therefore, the disclosure also provides industrial applications (*e.g.*, saccharification processes in ethanol production) using the enzymes and/or enzyme blends/compositions described herein. The enzymes and/or enzyme blends/compositions of the present disclosure can be used to decrease enzyme costs in biofuel production.

[0014] In one aspect, the disclosure pertains to enzymes (including variants thereof), or enzyme blends/compositions that are useful for hydrolyzing the major components of a lignocellulosic biomass (including, *e.g.*, cellulose, hemicellulose, and lignin) or any material comprising cellulose and/or hemicellulose. Such lignocellulosic biomass and/or material comprising cellulose and/or hemicellulose include, *e.g.*, seeds, grains, tubers, plant waste or byproducts of food processing or industrial processing (*e.g.*, stalks), corn (including, *e.g.*, cobs, stover, and the like), grasses (*e.g.*, Indian grass, such as *Sorghastrum nutans*; or, switchgrass, *e.g.*, *Panicum* species, such as *Panicum virgatum*), wood (including, *e.g.*, wood chips, processing waste), paper, pulp, recycled paper (*e.g.*, newspaper).

[0015] The enzyme blends/compositions can be used to hydrolyze cellulose comprising a linear chain of β -1,4-linked glucose moieties, or hemicellulose, of a complex structure that varies from plant to plant.

[0016] The enzyme blends/compositions can comprise a number of different enzymes, including, e.g., cellulases and/or hemicellulases. For example, the enzymes blends/compositions can be used to hydrolyze biomass or a suitable feedstock. The enzyme blends/compositions desirably comprise mixtures of enzymes, selected from, e.g., xylanases, xylosidases, cellobiohydrolases, arabinofuranosidases, and/or other enzymes that can digest hemicellulose to monomer sugars. An enzyme blend/composition can comprise a mixture of two or more, three or more, or four or more enzymes selected from one or more xylanases, one or more xylosidases, one or more cellobiohydrolases, one or more arabinofuranosidases, and one or more other enzymes that are capable of converting hemicellulose to monomer sugars. The other enzymes that can digest hemicellulose to monomer sugars include, without limitation, a cellulase, a hemicellulase, or a composition comprising a cellulase or a hemicellulase. An enzyme blend/composition can comprise a mixture of two or more, three or more, or four or more enzymes selected from a xylanase, a xylosidase, a cellobiohydrolase, an arabinofuranosidase, and at least one other enzyme capable of converting hemicellulose to monomer sugars. A non-limiting example of an enzyme blend/composition comprises a mixture of a xylanase, a xylosidase, a cellobiohydrolase, an arabinofuranosidase, and a β -glucosidase. The enzyme blend/composition is suitably one that is non-naturally occurring.

[0017] As used herein, the term "naturally occurring composition" refers to a composition produced by a naturally occurring source, which comprises one or more enzymatic components or activities, wherein each of these components or activities is found at the ratio and level produced by the naturally-occurring source as it is found in nature, untouched and unmodified by the human hand. Accordingly, a naturally occurring composition is, for example, one that is produced by an organism unmodified with respect to the cellulolytic or hemicellulolytic enzymes such that the ratio or levels of the component enzymes are unaltered from that produced by the native organism in its native environment. A "non-naturally occurring composition," on the other hand, refers to a composition produced by: (1) combining component cellulolytic or hemicellulolytic enzymes either in a naturally occurring ratio or a non-naturally occurring, i.e., altered, ratio; or (2) modifying an organism to express, overexpress or underexpress one or more endogenous or exogenous enzymes; or (3) modifying an organism such that at least one endogenous enzyme is deleted. A "non-naturally occurring composition" can also refer to a composition produced by a naturally-occurring and unmodified organism, but cultured in a man-made medium or environment that is different from the organism's native environment, such that the amounts or weight ratios of particular enzymes in the composition differ from those existing in a composition made by a native organism grown in its native habitat.

[0018] The enzyme blend/composition described herein is, for example, a fermentation broth. The fermentation broth can be one of a filamentous fungus, including, e.g., a *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, or *Chrysosporium*. An exemplary fungus of *Trichoderma* spp. is a *Trichoderma reesei*. An exemplary fungus of *Penicillium* spp. is a *Penicillium funiculosum*. The fermentation broth can be, for example, a cell-free fermentation

broth or a whole cell broth.

[0019] The enzyme blend/composition described herein is, in another example, a cellulase composition. The cellulase composition is a filamentous fungal cellulase composition, including, for example, a *Trichoderma*, such as a *Trichoderma reesei*, cellulase composition. The cellulase composition can be produced by a filamentous fungus, for example, a *Trichoderma*, such as a *Trichoderma reesei*.

[0020] For example, an enzyme blend/composition can comprise (a) one or more xylanase enzyme(s), wherein at least one of said one or more xylanase enzyme(s) is a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, or an AfuXyn5; (b) one or more β -xylosidase enzyme(s), wherein at least one of said one or more β -xylosidase enzyme(s) is a Group 1 β -xylosidase or a Group 2 β -xylosidase, wherein the Group 1 β -xylosidase is an Fv3A or an Fv43A, and the Group 2 β -xylosidase is a Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fo43A, an Fv43B, a Pa51A, a Gz43A, or a *Trichoderma reesei* Bxl1; (c) one or more L- α -arabinofuranosidase enzyme(s), wherein at least one of said one or more L- α -arabinofuranosidase is an Af43A, an Fv43B, a Pf51A, a Pa51A, or an Fv51A; (d) one or more cellulase enzymes; and optionally (e) one or more other components. The enzyme blend/composition is suitably one that is non-naturally occurring. The one or more cellulase enzyme(s) of (d) is desirably able to achieve at least 0.00005 fraction product per mg protein per gram of phosphoric acid swollen cellulose (PASC) as determined by a calcofluor assay. In a non-limiting example, the combined weight of xylanase enzyme(s) in the composition can represent or constitute 5 wt.% to 45 wt.% (e.g., 5 wt.% to 25 wt.%, 5 wt.% to 15 wt.%, 10 wt.% to 15 wt.%) of the combined or total protein weight in the composition, whereas the combined weight of β -xylosidase enzyme(s) can represent or constitute 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total proteins in the composition, whereas the combined weight of L- α -arabinofuranosidase enzyme(s) can represent or constitute 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 2 wt.% to 20 wt.%, 5 wt.% to 15 wt.%, 5 wt.% to 10 wt.%) of the combined or total protein weight in the composition, whereas the combined weight of cellulase enzyme(s) can represent or constitute 30 wt.% to 80 wt.% (e.g., 40 wt.% to 70 wt.%, 50 wt.% to 60 wt.%) of the combined or total protein weight in the composition. The enzyme blend/composition as described herein is, for example, a fermentation broth composition. The fermentation broth is, for example, one of a filamentous fungus, including, without limitation, a *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, or *Chrysosporium*. An exemplary fungus of *Trichoderma* spp. is a *Trichoderma reesei*. An exemplary fungus of *Penicillium* spp. is a *Penicillium funiculosum*. The fermentation can be, for example, a cell-free fermentation broth or a whole cell broth. The enzyme blend/composition as described herein can also be a cellulase composition, for example, a filamentous fungal cellulase composition. The cellulase composition, for example, can be produced by a filamentous fungus, such as by a *Trichoderma*.

[0021] For example, an enzyme blend/composition can comprise (a) one or more xylanase enzyme(s) wherein at least one of said one or more xylanase enzyme(s) is a *Trichoderma*

reesei Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, or an AfuXyn5; (b) one or both of Group 1 β -xylosidase enzymes: Fv3A and Fv43A; (c) one or more of Group 2 β -xylosidase enzyme(s) selected from Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fo43A, an Fv43B, a Pa51A, a Gz43A, and/or a *Trichoderma reesei* Bxl1; (d) one or more cellulase enzyme(s); and optionally (e) one or more other components. The one or more cellulase enzyme(s) of (e) is desirably able to achieve at least 0.00005 fraction product per mg protein per gram of phosphoric acid swollen cellulose (PASC) as determined by a calcofluor assay. The enzyme blend/composition is suitably one that is non-naturally occurring. In a non-limiting example, the combined weight of xylanase enzyme(s) in the composition can represent or constitute 5 wt.% to 45 wt.% (e.g., 5 wt.% to 25 wt.%, 5 wt.% to 15 wt.%, 10 wt.% to 15 wt.%) of the combined or total protein weight in the composition, whereas the combined weight of the Group 1 β -xylosidase enzyme(s) can constitute 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total protein weight in the composition, whereas the combined weight of the Group 2 β -xylosidase enzyme(s) can constitute 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total protein weight in the composition, and wherein the combined weight of the cellulase enzyme(s) can represent or constitute 30 wt.% to 80 wt.% (e.g., 40 wt.% to 70 wt.%, 50 wt.% to 60 wt.%) of the combined or total protein weight in the composition. The ratio of the weight of Group 1 β -xylosidase enzymes to the weight of Group 2 β -xylosidase enzymes can be, for example, 1:10 to 10:1 (e.g., 1:8 to 8:1, 1:6 to 6:1, 1:4 to 4:1, 1:2 to 2:1, or 1:1). The enzyme blend/composition can further comprise additional components, which may be accessory proteins or other protein/non-protein components. The additional components can constitute, for example, 1 wt.% to 50 wt.%, 1 wt.% to 10 wt.%, 2 wt.% to 5 wt.%, 5 wt.% to 10 wt.%, or 5 wt.% to 20 wt.% of the total weight of proteins in the composition. The enzyme blend/composition as described herein is, for example, a fermentation broth composition. The fermentation broth is, for example, one of a filamentous fungus, including, without limitation, a *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, or *Chrysosporium*. An exemplary fungus of *Trichoderma* spp. is a *Trichoderma reesei*. An exemplary fungus of *Penicillium* spp. is a *Penicillium funiculosum*. The fermentation can be, for example, a cell-free fermentation broth or a whole cell broth. The enzyme blend/composition as described herein can also be a cellulase composition, for example, a filamentous fungal cellulase composition. The cellulase composition, for example, can be produced by a filamentous fungus, such as by a *Trichoderma*.

[0022] In further examples, an enzyme blend/composition can comprise (a) one or more xylanase enzyme(s) wherein at least one of said one or more xylanase enzyme(s) is a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, or an AfuXyn5; (b) one or more β -xylosidase enzyme(s) wherein at least one of the one or more β -xylosidase enzyme(s) is a Group 1 β -xylosidase enzyme Fv3A or Fv43A or a Group 2 β -xylosidase enzyme Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fo43A, an Fv43B, a Pa51A, a Gz43A, and/or a *Trichoderma reesei* Bxl1; (c) one or more L- α -arabinofuranosidase enzyme(s), wherein at least one of said one or more L- α -arabinofuranosidase enzyme(s) is an Af43A, an Fv43B, a Pf51A, or an Fv51A; (d) one or more β -glucosidase enzyme(s); and optionally (e) one or more other components. The enzyme blend/composition is suitably one that is non-naturally occurring. In

a non-limiting example, the combined weight of xylanase enzyme(s) in the composition can represent or constitute 5 wt.% to 45 wt.% (e.g., 5 wt.% to 25 wt.%, 5 wt.% to 15 wt.%, 10 wt.% to 15 wt.%) of the combined or total protein weight in the composition, whereas the combined weight of the β -xylosidase enzyme(s) can constitute 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total protein weight in the composition, whereas the combined weight of the L- α -arabinofuranosidase enzyme(s) can constitute 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total protein weight in the composition, and wherein the combined weight of the β -glucosidase enzyme(s) can constitute 2 wt.% to 50 wt.% (e.g., up to 50 wt.%, 2 wt.% to 10 wt.%, or 3 wt.% to 8 wt.%) of the combined or total protein weight in the composition. The enzyme blend/composition can further comprise additional components, which may be accessory proteins or other protein/non-protein components. The additional components can constitute, for example, 1 wt.% to 50 wt.%, 1 wt.% to 10 wt.%, 2 wt.% to 5 wt.%, 5 wt.% to 10 wt.%, or 5 wt.% to 20 wt.% of the total weight of proteins in the composition. The enzyme blend/composition as described herein is, for example, a fermentation broth composition. The fermentation broth is, for example, one of a filamentous fungus, including, without limitation, a *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, or *Chrysosporium*. An exemplary fungus of *Trichoderma* spp. is a *Trichoderma reesei*. An exemplary fungus of *Penicillium* spp. is a *Penicillium funiculosum*. The fermentation can be, for example, a cell-free fermentation broth or a whole cell broth. The enzyme blend/composition as described herein can also be a cellulase composition, for example, a filamentous fungal cellulase composition. The cellulase composition, for example, can be produced by a filamentous fungus, such as by a *Trichoderma*.

[0023] An enzyme blend/composition can also comprise, for example, (a) one or more xylanase enzyme(s) wherein at least one of said one or more xylanase enzyme(s) is a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, or an AfuXyn5; (b) one or both of Group 1 β -xylosidase enzymes: Fv3A and Fv43A; (c) one or more of Group 2 β -xylosidase enzyme(s) selected from Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fo43A, an Fv43B, a Pa51A, a Gz43A, and/or a *Trichoderma reesei* Bxl1; and (d) one or more β -glucosidase enzyme(s); and optionally (e) one or more other components. The enzyme blend/composition is suitably one that is non-naturally occurring. In a non-limiting example, the combined weight of xylanase enzyme(s) constitutes 5 wt.% to 45 wt.% (e.g., 5 wt.% to 25 wt.%, 5 wt.% to 15 wt.%, 10 wt.% to 15 wt.%) of the total weight of proteins in the composition, whereas the combined weight of Group 1 β -xylosidase enzyme(s) constitutes 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total weight of proteins in the composition, whereas the combined weight of Group 2 β -xylosidase enzyme(s) constitutes 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total weight of proteins in the composition, whereas the combined weight of β -glucosidase enzyme(s) constitutes 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total weight of proteins in the composition. The ratio of the weight of Group 1 β -xylosidase enzymes to the weight of Group 2 β -xylosidase enzymes can be, for example, 1:10 to 10:1, for example, 1:8 to 8:1, 1:6 to 6:1, 1:4 to 4:1, 1:2 to 2:1, or 1:1. The enzyme blend/composition can further comprise additional components, which may be accessory

proteins or other protein/non-protein components. The additional components can constitute, for example, 1 wt.% to 50 wt.%, 1 wt.% to 10 wt.%, 2 wt.% to 5 wt.%, 5 wt.% to 10 wt.%, or 5 wt.% to 20 wt.% of the total weight of proteins in the composition. The enzyme blend/composition as described herein is, for example, a fermentation broth composition. The fermentation broth is, for example, one of a filamentous fungus, including, without limitation, a *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, or *Chrysosporium*. An exemplary fungus of *Trichoderma* spp. is a *Trichoderma reesei*. An exemplary fungus of *Penicillium* spp. is a *Penicillium funiculosum*. The fermentation can be, for example, a cell-free fermentation broth or a whole cell broth. The enzyme blend/composition as described herein can also be a cellulase composition, for example, a filamentous fungal cellulase composition. The cellulase composition, for example, can be produced by a filamentous fungus, such as by a *Trichoderma*.

[0024] Moreover, an enzyme blend/composition can comprise, for example, (a) one or more xylanase enzyme(s) wherein at least one of said one or more xylanase enzyme(s) is a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, or an AfuXyn5; (b) one or more β -xylosidase enzyme(s) wherein at least one of said one or more β -xylosidase enzyme(s) is a Group 1 β -xylosidase or a Group 2 β -xylosidase, wherein Group 1 β -xylosidase can be an Fv3A or an Fv43A, and a Group 2 β -xylosidase can be a Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fo43A, an Fv43B, a Pa51A, a Gz43A, and/or a *Trichoderma reesei* Bxl1; and (c) one or more L- α -arabinofuranosidase enzyme(s), wherein at least one of said one or more L- α -arabinofuranosidase enzyme(s) is an Af43A, an Fv43B, a Pf51A, or an Fv51A; and optionally (d) one or more other components. The enzyme blend/composition is suitably one that is non-naturally occurring. In a non-limiting example, the combined weight of the xylanase enzyme(s) constitutes 5 wt.% to 45 wt.% (e.g., 5 wt.% to 25 wt.%, 5 wt.% to 15 wt.%, 10 wt.% to 15 wt.%) of the total protein weight in the composition, whereas the combined weight of the β -xylosidase enzyme(s) constitutes 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total protein weight of the composition, whereas the combined weight of L- α -arabinofuranosidase enzyme(s) constitutes 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total protein weight of the composition. The enzyme blend/composition can further comprise additional components, which may be accessory proteins or other protein/non-protein components. The additional components can constitute, for example, 1 wt.% to 50 wt.%, 1 wt.% to 10 wt.%, 2 wt.% to 5 wt.%, 5 wt.% to 10 wt.%, or 5 wt.% to 20 wt.% of the total weight of proteins in the composition. The enzyme blend/composition as described herein is, for example, a fermentation broth composition. The fermentation broth is, for example, one of a filamentous fungus, including, without limitation, a *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, or *Chrysosporium*. An exemplary fungus of *Trichoderma* spp. is a *Trichoderma reesei*. An exemplary fungus of *Penicillium* spp. is a *Penicillium funiculosum*. The fermentation can be, for example, a cell-free fermentation broth or a whole cell broth. The enzyme blend/composition as described herein can also be a cellulase composition, for example, a filamentous fungal cellulase composition. The cellulase composition, for example, can be produced by a filamentous fungus, such as by

a *Trichoderma*.

[0025] An enzyme blend/composition can also be one that comprises (a) one or more xylanase enzyme(s) wherein at least one of said one or more xylanase enzyme(s) is a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, or an AfuXyn5; (b) one or both of Group 1 β -xylosidase enzymes: Fv3A and Fv43A; (c) one or more of Group 2 β -xylosidase enzyme(s) selected from Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fo43A, an Fv43B, a Pa51A, a Gz43A, and/or a *Trichoderma reesei* Bxl1; and optionally (d) one or more other components. The enzyme blend/composition is suitably one that is non-naturally occurring. In a non-limiting example, the combined weight of xylanase enzyme(s) can constitute 5 wt.% to 45 wt.% (e.g., 5 wt.% to 25 wt.%, 5 wt.% to 15 wt.%, 10 wt.% to 15 wt.%) of the total protein weight in the composition, whereas the combined weight of Group 1 β -xylosidase enzyme(s) can constitute 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total protein weight in the composition, whereas the combined weight of Group 2 β -xylosidase enzyme(s) can constitute 2 wt.% to 50 wt.% (e.g., 2 wt.% to 30 wt.%, 5 wt.% to 25 wt.%, 5 wt.% to 10 wt.%) of the total protein weight in the composition. The ratio of the weight of Group 1 β -xylosidase enzymes to the weight of Group 2 β -xylosidase enzymes can be, for example, 1:10 to 10:1, for example, 1:8 to 8:1, 1:6 to 6:1, 1:4 to 4:1, 1:2 to 2:1, or 1:1. The enzyme blend/composition can further comprise additional components, which may be accessory proteins or other protein/non-protein components. The additional components can constitute, for example, 1 wt.% to 50 wt.%, 1 wt.% to 10 wt.%, 2 wt.% to 5 wt.%, 5 wt.% to 10 wt.%, or 5 wt.% to 20 wt.% of the total weight of proteins in the composition. The enzyme blend/composition as described herein is, for example, a fermentation broth composition. The fermentation broth is, for example, one of a filamentous fungus, including, without limitation, a *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, or *Chrysosporium*. An exemplary fungus of *Trichoderma* spp. is a *Trichoderma reesei*. An exemplary fungus of *Penicillium* spp. is a *Penicillium funiculosum*. The fermentation can be, for example, a cell-free fermentation broth or a whole cell broth. The enzyme blend/composition as described herein can also be a cellulase composition, for example, a filamentous fungal cellulase composition. The cellulase composition, for example, can be produced by a filamentous fungus, such as by a *Trichoderma*.

[0026] The enzymes, enzyme blends/compositions of the disclosure can be used in the food industry, e.g., for baking, for fruit and vegetable processing, in breaking down of agricultural waste, in the manufacture of animal feed, in pulp and paper production, in textile manufacture, or in household and industrial cleaning agents. The enzymes, and the enzymes in the enzyme blends/compositions of the disclosure are, for example, each independently produced by a microorganism, e.g., by a fungi or a bacteria.

[0027] The enzymes, enzyme blends/compositions of the disclosure can also be used as commercial enzymes or compositions to digest lignocellulose from any suitable sources, including all biological sources, such as plant biomasses, e.g., corn, grains, grasses (e.g., Indian grass, such as *Sorghastrum nutans*; or, switchgrass, e.g., *Panicum* species, such as

Panicum virgatum), or, woods or wood processing byproducts, e.g., in the wood processing, pulp and/or paper industry, in textile manufacture, in household and industrial cleaning agents, and/or in biomass waste processing.

[0028] In another aspect, the disclosure provides isolated, synthetic or recombinant nucleic acids having at least about 70%, for example, at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%; 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete (100%) sequence identity to a nucleic acid sequence of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 46, 47, 48, 49, or 50, over a region of at least about 10, e.g., at least about 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, or 2000, residues. Relatedly, the disclosure provides isolated, synthetic, or recombinant nucleic acids that are capable of hybridizing, under high stringency conditions, to a complement of 97%, 98%, 99%, or complete (100%) sequence identity to a nucleic acid sequence of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 46, 47, 48, 49, or 50, or to a fragment thereof. The fragment, for example, can be at least 150 contiguous residues in length, for example, at least 200, 250, or 300 contiguous residues in length. The present disclosure provides nucleic acids encoding a polypeptide having hemicellulolytic activity. Exemplary hemicellulolytic activity includes, without limitation, xylanase, β -xylosidase, and/or L- α -arabinofuranosidase activity. Exemplary polypeptides having hemicellulolytic activity include, without limitation, a xylanase, a β -xylosidase, and/or an L- α -arabinofuranosidase.

[0029] The disclosure further provides isolated, synthetic, or recombinant nucleic acids encoding an enzyme of the disclosure, including a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, or 44, or a subsequence thereof (e.g., a catalytic domain ("CD") or carbohydrate binding module ("CBM")), or a suitable variant thereof. In some embodiments, a nucleic acid of the disclosure encodes the mature portion of a protein of amino acid sequence SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, or 44, which is optionally operably linked to a heterologous signal sequence, e.g., the *Trichoderma reesei* CBHI signal sequence. The nucleic acid desirably encodes a polypeptide having hemicellulolytic activity, e.g., xylanase, β -xylosidase, and/or L- α -arabinofuranosidase activity. The nucleic acid of the disclosure encodes a hemicellulase, for example, a xylanase, a β -xylosidase, and/or an L- α -arabinofuranosidase, or a suitable variant thereof. Further nucleic acids of the disclosure are described in Section 6.2 below.

[0030] The disclosure additionally provides expression cassettes comprising a nucleic acid of the disclosure or a subsequence thereof. For example, the nucleic acid comprises at least about 70%, e.g., at least about 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to a nucleic acid sequence of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 46, 47, 48, 49, or 50, over a region of at least about 10 residues, e.g., at least about 10, 20, 30, 40, 50, 75, 90, 100, 150, 200, 250,

300, 350, 400, or 500 residues. In another example, the nucleic acid encodes a polypeptide of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, or 44, wherein the nucleic acid is optionally operably linked to a promoter. The promoter can be, e.g., a fungal, viral, bacterial, mammalian, or plant promoter. The promoter can be a constitutive promoter or an inducible promoter. An exemplary suitable promoter is expressible in filamentous fungi, e.g., *Trichoderma reesei*. A suitable promoter can be derived from a filamentous fungus, e.g., *Trichoderma reesei*, e.g., a cellobiohydrolase I ("*cbh1*") gene promoter from *Trichoderma reesei*.

[0031] The disclosure further provides a recombinant cell engineered to express a nucleic acid of the disclosure or an expression cassette of the disclosure. For example, the nucleic acid comprises at least about 70%, e.g., at least about 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to a nucleic acid sequence of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 46, 47, 48, 49, or 50, over a region of at least about 10, e.g., at least about 15, 20, 25, 50, 75, 100, 150, 200, 250, 300, 400, or 500 nucleotide residues. The nucleic acid can encode a polypeptide of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, or 44, wherein the nucleic acid is optionally operably linked to a promoter. The expression cassette can comprise the nucleic acid having at least about 70% (e.g., at least about 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%) sequence identity to SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 46, 47, 48, 49, or 50, over a region of at least about 10 nucleotide residues. For example, the expression cassette can comprise the nucleic acid encoding a polypeptide of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, or 44, wherein the nucleic acid is optionally operably linked to a promoter. The recombinant cell is desirably a recombinant bacterial cell, a recombinant mammalian cell, a recombinant fungal cell, a recombinant yeast cell, a recombinant insect cell, a recombinant algal cell, or a recombinant plant cell. For example, the recombinant cell is a recombinant filamentous fungal cell, such as a *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, or *Chrysosporium* cell.

[0032] The disclosure provides transgenic plants comprising a nucleic acid of the disclosure or an expression cassette of the disclosure.

[0033] The disclosure provides isolated, synthetic or recombinant polypeptides comprising an amino acid sequence having at least about 80%, e.g., at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%, or complete (100%) sequence identity to a polypeptide of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, or 44, over a region of at least about 10, e.g., at least about 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, or 350 residues, or over the full length immature polypeptide, the full length mature polypeptide, the full length CD, or the full length CBM. Exemplary polypeptides include fragments of at least about 10, for example, at least about 15, 20, 25, 30, 35, 40, 45, 50, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500,

550, 600 residues in length. In specific embodiments, the fragments comprise a CD and/or a CBM. Where a fragment comprises both a CD and a CBM of an enzyme of the disclosure, the fragment optionally includes a linker separating the CD and the CBM. The linker can be a native linker or a heterologous linker. In certain embodiments, the polypeptides of the disclosure have one or more hemicellulase activities. Polypeptides or peptide sequences of the disclosure include sequences encoded by the nucleic acids of the disclosure. Exemplary polypeptides are described in Section 6.1.

[0034] The disclosure additionally provides a chimeric or fusion protein comprising at least one domain of a polypeptide of the disclosure (*e.g.*, the CD, the CBM, or both). The at least one domain can be operably linked to a second amino acid sequence, *e.g.*, a signal peptide sequence.

[0035] Conversely, the disclosure provides a chimeric or fusion protein comprising a signal sequence of a polypeptide of the disclosure operably linked to a second sequence, *e.g.*, encoding the amino acid sequence of a heterologous polypeptide that is not naturally associated with the signal sequence. Accordingly, the disclosure provides a recombinant polypeptide comprising residues 1 to 13, 1 to 14, 1 to 15, 1 to 16, 1 to 17, 1 to 18, 1 to 19, 1 to 20, 1 to 21, 1 to 22, 1 to 23, 1 to 24, 1 to 25, 1 to 26, 1 to 27, 1 to 28, 1 to 28, 1 to 30, 1 to 31, 1 to 32, 1 to 33, 1 to 34, 1 to 35, 1 to 36, 1 to 37, 1 to 38, or 1 to 40 of a polypeptide of, *e.g.*, SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, or 44. Further exemplary chimeric or fusion polypeptides are described in Section 6.1.1.

[0036] The disclosure also provides methods of producing a recombinant polypeptide comprising: (a) culturing a host cell engineered to express a polypeptide of the disclosure; and (b) recovering the polypeptide. The recovery of the polypeptide includes, *e.g.*, recovery of the fermentation broth comprising the polypeptide. In certain embodiments, recovery of the polypeptide can include further purification step(s).

[0037] The disclosure provides methods for hydrolyzing, breaking up, or disrupting a cellooligosaccharide, an arabinoxylan oligomer, or a glucan- or cellulose-comprising composition comprising contacting the composition with an enzyme, enzyme blend/composition of the disclosure under suitable conditions, wherein the enzyme, or enzyme blend/composition hydrolyzes, breaks up, or disrupts the cellooligosaccharide, arabinoxylan oligomer, or glucan- or cellulose-comprising composition.

[0038] The disclosure provides enzyme "blends" or compositions (also termed "enzyme blend/composition" herein), comprising a polypeptide of the disclosure, or a polypeptide encoded by a nucleic acid of the disclosure. In some embodiments, the polypeptide of the disclosure has one or more activities selected from xylanase, β -xylosidase, and L- α -arabinofuranosidase activities. In certain embodiments, the enzyme blends/compositions are used or are useful for depolymerization of cellulosic and hemicellulosic polymers to metabolizable carbon moieties. The enzyme blends of the disclosure can be in the form of a composition *e.g.*, as a product of manufacture. The composition can be, *e.g.*, a formulation,

and can take the physical form of, e.g., a liquid or a solid. In exemplary embodiments, an enzyme blend/composition of the disclosure includes a cellulase comprising at least three different enzyme types selected from (1) an endoglucanase, (2) a cellobiohydrolase, and (3) a β -glucosidase; or at least three different enzymatic activities selected from (1) an endoglucanase activity catalyzing the cleavage of internal β -1,4 linkages of cellulosic or hemicellulosic materials, resulting in shorter glucooligosaccharides, (2) a cellobiohydrolase activity catalyzing the cleavage and release, in an "exo" manner, of cellobiose units (e.g., β -1,4 glucose-glucose disaccharide), and (3) a β -glucosidase activity catalyzing the release of glucose monomers from short cellooligosaccharides (e.g., cellobiose). Exemplary enzyme blends/compositions of the disclosure are described in Section 6.3.4. below.

[0039] In another aspect, the disclosure provides methods for processing a biomass material comprising lignocellulose comprising contacting a composition comprising a cellulose and/or a fermentable sugar with a polypeptide of the disclosure, or a polypeptide encoded by a nucleic acid of the disclosure, or an enzyme blend/composition (e.g., a product of manufacture) of the disclosure. Suitable biomass material comprising lignocellulose can be derived from an agricultural crop, a byproduct of a food or feed production, a lignocellulosic waste product, a plant residue, or a waste paper or waste paper product. The polypeptides of the disclosure can have one or more enzymatic activities selected from cellulase, endoglucanase, cellobiohydrolase, β -glucosidase, xylanase, mannanase, β -xylosidase, arabinofuranosidase, and other hemicellulase activities. Suitable plant residue can comprise grain, seeds, stems, leaves, hulls, husks, corncobs, corn stover, straw, grasses, wood, wood chips, wood pulp and sawdust. The grasses can be, e.g., Indian grass, or switchgrass. The grasses can also be, for example, *Miscanthus*. The paper waste can be, e.g., discarded or used photocopy paper, computer printer paper, notebook paper, notepad paper, typewriter paper, newspapers, magazines, cardboard, and various paper-based packaging materials. The paper waste can also be, for example, pulp.

[0040] The disclosure provides compositions (including enzymes, enzyme blends/compositions, e.g., products of manufacture of the disclosure) comprising a mixture of hemicellulose- and cellulose-hydrolyzing enzymes, and at least one biomass material. Optionally the biomass material comprises a lignocellulosic material derived from an agricultural crop. Alternatively the biomass material is a byproduct of a food or feed production. Suitable biomass material can also be a lignocellulosic waste product, a plant residue, a waste paper or waste paper product, or comprises a plant residue. The plant residue can, e.g., be one comprising grains, seeds, stems, leaves, hulls, husks, corncobs, corn stover, grasses, straw, wood, wood chips, wood pulp, or sawdust. Exemplary grasses include, without limitation, Indian grass or switchgrass. Exemplary grasses can also include *Miscanthus*. Exemplary paper waste include, without limitation, discarded or used photocopy paper, computer printer paper, notebook paper, notepad paper, typewriter paper, newspapers, magazines, cardboard and paper-based packaging materials. Exemplary paper waste can also include, e.g., pulp.

5. BRIEF DESCRIPTION OF THE FIGURES AND TABLES

[0041]

Tables 1A-1B: Table 1A provides a summary of the sequence identifiers used in the present disclosure for glycosyl hydrolase enzymes; Table 1B provides accession numbers for additional glycosyl hydrolase enzymes referred to in the Examples.

Table 2: Activity of expressed proteins on synthetic substrates pNPA and pNPX (as defined in Section 7.1.6. below) in terms relative to the *Trichoderma reesei* Quad delete host background activity. The Quad delete host background (or "XQuad") is defined as the activity of the expressed protein(s) in the Quad delete strain divided by the activity of the Quad delete background strain without the expressed protein(s). For example, a value of >1 indicates that the expressed protein have an activity greater than that of the background.

Table 3: Xylanase activity of purified candidate endo-xylanases with birchwood xylan, incubated at 50°C, pH 5.

Table 4: Percent conversion of cob arabinoxylan oligomers to monomer products based on total sugar available as determined by acid hydrolysis. See, Example 4.

Table 5: Experiment results (as described in Example 7) defining the level of hemicellulase activity for hydrolyzing treated corncob to monomer sugars. The column entitled "run #" indicates the randomized experimental order. The column entitled "trial #" indicates the standard experimental design order. The column entitled "Quad" indicates fraction of the total protein that is from the culture supernatant for a growth of the Quad deleted *T. reesei* strain. The column entitled "Xyn3" indicates fraction of the total protein that is *Trichoderma reesei* Xyn3. The column entitled "Fv43D" indicates fraction of the total protein that is Fv43D. The column entitled "Fv51 A" indicates fraction of the total protein that is Fv51A. The column entitled "Fv43A" indicates fraction of the total protein that is Fv43A. The column entitled "Fv43B" indicates fraction of the total protein that is Fv43B. The column entitled "loading (ug/mg carbo)" indicates the protein loaded into the saccharification reaction in units of micrograms of protein per milligram of carbohydrate. The columns entitled "Xyl mg/mL, Glu mg/mL, Arab mg/mL, and G+X+A mg/mL" indicate the concentration of xylose, glucose, arabinose, and the combination of those three sugar products that is detected at the end of the saccharification reaction. The columns entitled "Xyl %theor, Glu %theor, and Arab %theor" indicate the percent of theoretical yield of xylose, glucose, and arabinose reached at the end of the saccharification reaction.

Table 6: Calculated ratios of the seven enzymatic components for predicted maximal yield of glucose, xylose and arabinose from hydrolysis of corncob. The column "loading total mg/gr carb" indicates the total enzyme dose at which the predictions are calculated. The rows entitled "Total mg/ml G+X+A, % Yield Glucose, % Yield Xylose, and % Yield Arabinose" indicate the response for which the optimum has been calculated. The column entitled "r2 data fit to model (includes both loadings)" indicates the r-squared statistical parameter for the model fit to data presented in Table 5. The columns entitled "fraction Accellerase, fraction Quad del sup, fraction purified *Trichoderma reesei* Xyn3, fraction purified Fv43D, fraction purified Fv51A,

fraction purified Fv43A, and fraction purified Fv43B" indicate the fraction of that component that is calculated to be optimal by the model fitted to the data in Table 5.

Table 7: Refinement of enzyme loadings for maximal hydrolytic conversion of corncob using blends including Fv3A and Fv43D enzymes in 1.06 g, 14% dry solids pretreated cob reactions. The conditions for saccharification were as described in Example 7. Columns marked with enzyme names indicate the mg of each of the listed enzyme per gram of glucan or xylan used.

Table 8: Refinement of enzyme loadings for maximal conversion of corncob using blends containing Fv51A as the only L- α -arabinofuranosidase in 1.06 gr, 14% dry solids pretreated cob reactions. The conditions used for saccharification were as described in Example 7. Columns marked with enzyme names indicate the mg of each of the listed enzymes per gram of glucan or xylan used. Columns marked with carbohydrates indicate the mg per mL of each carbohydrate product produced based on measurements made with size exclusion chromatography. The >dp2 column includes all oligomers larger than a disaccharide.

Table 9: Sugar yields obtained from mixes A, B, C of purified hemicellulases tested in 1.06 g, 14% dry solids pretreated cob reactions. The conditions used for saccharification were as described in Example 7. Mix A: 6 mg *Trichoderma reesei* Xyn3, 4 mg Fv3A, 1 mg Fv51A per gram xylan. Mix B: 6 mg *Trichoderma reesei* Xyn3, 1 mg Fv43D, 3 mg Fv43A, 3 mg Fv43B per gram xylan. Mix C: 6 mg *Trichoderma reesei* Xyn3, 3 mg Fv3A, 1 mg Fv43D, 1 mg Fv51A per gram xylan. Columns marked with monomer sugars indicate the % yield of each of the listed monomer.

Table 10: Sugar yields from treatment of hemicellulose preparations made from corncob, sorghum, switchgrass and sugar cane bagasse by hemicellulase mixes A, B, C. The reactions were run at 100- μ L scale in 50 mM pH 5.0 Sodium Acetate buffer for 6 h at 48°C as described in Example 8. The % yield for each monomer sugar is shown.

Table 11: Concentration of the majority enzymes expressed by various *T. reesei* integrated expression strains (designated H3A, 39A, 69A, A10A, G6A, 102, 44A, 11A, G9A) as determined by percent of the integrated HPLC area.

Table 12: List of switchgrass pretreatment parameters and saccharification results from the various pretreatments.

Table 13: Saccharification results of a hardwood pulp with the enzyme composition produced by a *T. reesei* integrated strain, H3A, in reactions with different amounts of solids, enzymes, and incubation time.

Table 14: Saccharification of a hardwood pulp with the enzyme composition produced by an integrated strain H3A over a temperature range and a pH range.

Signal sequences listed below and in the figures were predicted. The predictions were made with the SignalP algorithm (*available at* <http://www.cbs.dtu.dk>). Domain predictions were made based on one or more of the Pfam, SMART, or NCBI databases.

Figures 1A-1B: Figure 1A: Fv3A nucleotide sequence (SEQ ID NO:1). Figure 1B: Fv3A amino

acid sequence (SEQ ID NO:2). SEQ ID NO:2 is the sequence of the immature Fv3A. Fv3A has a predicted signal sequence corresponding to positions 1 to 23 of SEQ ID NO:2 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 24 to 766 of SEQ ID NO:2. The predicted conserved domain is in boldface type.

Figures 2A-2B: Figure 2A: Pf43A nucleotide sequence (SEQ ID NO:3). Figure 2B: Pf43A amino acid sequence (SEQ ID NO:4). SEQ ID NO:4 is the sequence of the immature Pf43A. Pf43A has a predicted signal sequence corresponding to positions 1 to 20 of SEQ ID NO:4 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 21 to 445 of SEQ ID NO:4. The predicted conserved domain is in boldface type, the predicted carbohydrate binding module ("CBM") is in uppercase type, and the predicted linker separating the CD and CBM is in italics.

Figures 3A-3B: Figure 3A: Fv43E nucleotide sequence (SEQ ID NO:5). Figure 3B: Fv43E amino acid sequence (SEQ ID NO:6). SEQ ID NO:6 is the sequence of the immature Fv43E. Fv43E has a predicted signal sequence corresponding to positions 1 to 18 of SEQ ID NO:6 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 19 to 530 of SEQ ID NO:6. The predicted conserved domain is in boldface type.

Figures 4A-4B: Figure 4A: Fv39A nucleotide sequence (SEQ ID NO:7). Figure 4B: Fv39A amino acid sequence (SEQ ID NO:8). SEQ ID NO:8 is the sequence of the immature Fv39A. Fv39A has a predicted signal sequence corresponding to positions 1 to 19 of SEQ ID NO:8 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 20 to 439 of SEQ ID NO:8. The predicted conserved domain is in boldface type.

Figures 5A-5B: Figure 5A: Fv43A nucleotide sequence (SEQ ID NO:9). Figure 5B: Fv43A amino acid sequence (SEQ ID NO:10). SEQ ID NO:10 is the sequence of the immature Fv43A. Fv43A has a predicted signal sequence corresponding to positions 1 to 22 of SEQ ID NO:10 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 23 to 449 of SEQ ID NO:10. The predicted conserved domain is in boldface type, the predicted CBM is in uppercase type, and the predicted linker separating the conserved domain and CBM is in italics.

Figures 6A-6B: Figure 6A: Fv43B nucleotide sequence (SEQ ID NO:11). Figure 6B: 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 positions 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 positions 17 to 574 of SEQ ID NO:12. The predicted conserved domain is in boldface type.

Figures 7A-7B: Figure 7A: Pa51A nucleotide sequence (SEQ ID NO:13). Figure 7B: Pa51A amino acid sequence (SEQ ID NO:14). SEQ ID NO:14 is the sequence of the immature Pa51A. Pa51A has a predicted signal sequence corresponding to positions 1 to 20 of SEQ ID NO:14

(underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 21 to 676 of SEQ ID NO:14. The predicted L- α -arabinofuranosidase conserved domain is in boldface type. For expression purposes, the genomic DNA was codon optimized for expression in *T. reesei* (see Figure 60B).

Figures 8A-8B: Figure 8A: Gz43A nucleotide sequence (SEQ ID NO:15). Figure 8B: Gz43A amino acid sequence (SEQ ID NO:16). SEQ ID NO:16 is the sequence of the immature Gz43A. Gz43A has a predicted signal sequence corresponding to positions 1 to 18 of SEQ ID NO:16 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 19 to 340 of SEQ ID NO:16. The predicted conserved domain is in boldface type. For expression purposes, the Gz43A predicted signal sequence was replaced by the *T. reesei*CBH1 signal sequence (myrklavisaflatara (SEQ ID NO: 51)) in *T. reesei* (see Figure 61).

Figures 9A-9B: Figure 9A: Fo43A nucleotide sequence (SEQ ID NO:17). Figure 9B: Fo43A amino acid sequence (SEQ ID NO:18). SEQ ID NO:18 is the sequence of the immature Fo43A. Fo43A has a predicted signal sequence corresponding to positions 1 to 20 of SEQ ID NO:18 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 21 to 348 of SEQ ID NO:18. The predicted conserved domain is in boldface type. For expression purposes, the Fo43A predicted signal sequence was replaced by the *T. reesei*CBH1 signal sequence (myrklavisaflatara (SEQ ID NO:51)) (see Figure 62).

Figures 10A-10B: Figure 10A: Af43A nucleotide sequence (SEQ ID NO:19). Figure 10B: Af43A amino acid sequence (SEQ ID NO:20). SEQ ID NO:20 is the sequence of the immature Af43A. The predicted conserved domain is in boldface type.

Figures 11A-11B: Figure 11A: Pf51A nucleotide sequence (SEQ ID NO:21). Figure 11B: Pf51A amino acid sequence (SEQ ID NO:22). SEQ ID NO:22 is the sequence of the immature Pf51A. Pf51A has a predicted signal sequence corresponding to positions 1 to 20 of SEQ ID NO:22 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 21 to 642 of SEQ ID NO:22. The predicted L- α -arabinofuranosidase conserved domain is in boldface type. For expression purposes, the predicted Pf51A signal sequence was replaced by a codon optimized the *T. reesei* CBH1 signal sequence (myrklavisaflatara (SEQ ID NO:51)) (underlined) and the Pf51A nucleotide sequence was codon optimized for expression in *T. reesei* (see Figure 63).

Figures 12A-12B: Figure 12A: AfuXyn2 nucleotide sequence (SEQ ID NO:23). Figure 12B: AfuXyn2 amino acid sequence (SEQ ID NO:24). SEQ ID NO:24 is the sequence of the immature AfuXyn2. AfuXyn2 has a predicted signal sequence corresponding to positions 1 to 18 of SEQ ID NO:24 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 19 to 228 of SEQ ID NO:24. The predicted GH11 conserved domain is in boldface type.

Figures 13A-13B: Figure 13A: AfuXyn5 nucleotide sequence (SEQ ID NO:25). Figure 13B: AfuXyn5 amino acid sequence (SEQ ID NO:26). SEQ ID NO:26 is the sequence of the

immature AfuXyn5. AfuXyn5 has a predicted signal sequence corresponding to positions 1 to 19 of SEQ ID NO:26 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 20 to 313 of SEQ ID NO:26. The predicted GH11 conserved domain is in boldface type.

Figures 14A-14B: Figure 14A: Fv43D nucleotide sequence (SEQ ID NO:27). Figure 14B: Fv43D amino acid sequence (SEQ ID NO:28). SEQ ID NO:28 is the sequence of the immature Fv43D. Fv43D has a predicted signal sequence corresponding to positions 1 to 20 of SEQ ID NO:28 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 21 to 350 of SEQ ID NO:28. The predicted conserved domain is in boldface type.

Figures 15A-15B: Figure 15A: Pf43B nucleotide sequence (SEQ ID NO:29). Figure 15B: Pf43B amino acid sequence (SEQ ID NO:30). SEQ ID NO:30 is the sequence of the immature Pf43B. Pf43B has a predicted signal sequence corresponding to positions 1 to 20 of SEQ ID NO:30 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 21 to 321 of SEQ ID NO:30. The predicted conserved domain is in boldface type.

Figures 16A-16B: Figure 16A: Fv51A nucleotide sequence (SEQ ID NO:31). Figure 16B: Fv51A amino acid sequence (SEQ ID NO:32). SEQ ID NO:32 is the sequence of the immature Fv51A. Fv51A has a predicted signal sequence corresponding to positions 1 to 19 of SEQ ID NO:32 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 20 to 660 of SEQ ID NO:32. The predicted L- α -arabinofuranosidase conserved domain is in boldface type.

Figures 17A-17B: Figure 17A: Cg51B nucleotide sequence (SEQ ID NO:33). Figure 17B: Cg51B amino acid sequence (SEQ ID NO:34). SEQ ID NO:34 is the sequence of the immature Cg51B. Cg51B has a predicted signal sequence corresponding to positions 1 to 20 of SEQ ID NO:34 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 21 to 670 of SEQ ID NO:34. The predicted conserved domain is in boldface type.

Figures 18A-18B: Figure 18A: Fv43C nucleotide sequence (SEQ ID NO:35). Figure 18B: Fv43C amino acid sequence (SEQ ID NO:36). SEQ ID NO:36 is the sequence of the immature Fv43C. Fv43C has a predicted signal sequence corresponding to positions 1 to 22 of SEQ ID NO:36 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 23 to 333 of SEQ ID NO:36. The predicted conserved domain is in boldface type.

Figures 19A-19B: Figure 19A: Fv30A nucleotide sequence (SEQ ID NO:37). Figure 19B: Fv30A amino acid sequence (SEQ ID NO:38). SEQ ID NO:38 is the sequence of the immature Fv30A. Fv30A has a predicted signal sequence corresponding to positions 1 to 19 of SEQ ID NO:38 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 20 to 537 of SEQ ID NO:38.

Figures 20A-20B: Figure 20A: Fv43F nucleotide sequence (SEQ ID NO:39). Figure 20B: Fv43F amino acid sequence (SEQ ID NO:40). SEQ ID NO:40 is the sequence of the immature Fv43F. Fv43F has a predicted signal sequence corresponding to positions 1 to 20 of SEQ ID NO:40 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 21 to 315 of SEQ ID NO:40.

Figures 21A-21B: Figure 21 A: *Trichoderma reesei* Xyn3 nucleotide sequence (SEQ ID NO:41). Figure 21B: *Trichoderma reesei*Xyn3 amino acid sequence (SEQ ID NO:42). SEQ ID NO:42 is the sequence of the immature *Trichoderma reesei*Xyn3. *Trichoderma reesei* Xyn3 has a predicted signal sequence corresponding to positions 1 to 16 of SEQ ID NO:42 (underlined); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to positions 17 to 347 of SEQ ID NO:42. The predicted conserved domain is in bold face type.

Figure 22: Amino acid sequence of *Trichoderma reesei* Xyn2 (SEQ ID NO:43). The signal sequence is underlined. The predicted conserved domain is in bold face type. The coding sequence can be found in Törrönen et al. Biotechnology, 1992, 10:1461-65.

Figure 23: Amino acid sequence of *Trichoderma reesei* Bxl1 (SEQ ID NO:44). The signal sequence is underlined. The predicted conserved domain is in bold face type. The coding sequence can be found in Margolles-Clark et al. Appl. Environ. Microbiol. 1996, 62(10):3840-46.

Figure 24: Amino acid sequence of *Trichoderma reesei* Bgl1 (SEQ ID NO:45). The signal sequence is underlined. The predicted conserved domain is in bold face type. The coding sequence can be found in Barnett et al. Bio-Technology, 1991, 9(6) :562-567.

Figure 25: Cellulase activity assay using PASC hydrolysis with calcofluor detection.

Figure 26: Xylanase elution profile.

Figure 27: SDS-PAGE detection of the two step separation of AfuXyn 5. Lane 1: Crude sample; Lane 2: Eluate from Phenyl column; Lane 3: Eluate from GF column.

Figure 28: pENTR/D-TOPO plasmid.

Figure 29: pTrex3gM.

Figures 30A-30B: Performance of different enzymes on corncob substrate. Error bars represent the experimental errors associated with triplicate cob assays. The numbers in parenthesis along the x-axis represent the enzyme doses in mg of protein per g of cellulose.

Figure 31: Performance of different enzyme blends/compositions on corncob substrate. Error bars represent the experimental errors associated with triplicate cob assays. The numbers in parentheses along the x-axis represent the enzyme doses in mg of protein per g cellulose.

Figure 32: Performance of different enzyme blends/compositions on corncob substrate. Error bars represent the experimental errors associated with triplicate cob assays. The numbers in

parentheses along the x-axis represent the enzyme doses in mg of protein per g cellulose.

Figure 33: Performance of different enzyme blends/compositions on corncob substrate. Error bars represent the experimental errors associated with triplicate cob assays. The numbers in parentheses along the x-axis represent the enzyme doses in mg of protein per g cellulose.

Figure 34: Performance of different enzyme blends/compositions on corncob substrate. Error bars represent the experimental errors associated with triplicate cob assays. The numbers in parentheses along the x-axis represent the enzyme doses in mg of protein per g cellulose.

Figures 35A-35C: Performance of different enzyme blends/compositions on corncob substrate. Error bars represent the experimental errors associated with triplicate cob assays. The numbers along the x-axis represent the enzyme doses in mg of protein per g cellulose.

Figure 36: Anomeric proton NMR region of short arabinoxylan oligomers shows cleavage by Fv43A plus Fv43B.

Figure 37: Anomeric proton NMR region of short arabinoxylan oligomers shows cleavage of β -1,2-linked xylose from arabinose by Fv3A.

Figure 38: Alignment between the amino acid sequences of *Trichoderma reesei* β -xylosidase and Fv3A.

Figure 39: pENTR-TOPO-Bgl1 (943/942) plasmid.

Figure 40: pTrex3g 943/942 Bgl1 expression vector.

Figure 41: pENTR- *Trichoderma reesei* Xyn3 plasmid.

Figure 42: pTrex3g/ *Trichoderma reesei* Xyn3 expression vector.

Figure 43: pENTR-Fv3A plasmid.

Figure 44: pTrex6g/Fv3A expression vector.

Figure 45: TOPO Blunt/Pgl1-Fv43D plasmid.

Figure 46: TOPO Blunt/Pgl1-Fv51A plasmid.

Figures 47A-47D: Glucan (Figure 47A) and xylan (Figure 47B) conversions to monomer sugars by secreted enzyme fermentation broths from *T. reesei* integrated expression strains. The 3-day sample was analyzed for the extent of conversion of glucan and xylan to both monomer and soluble oligomer products (Figure 47C). Figure 47D shows a chromatographic comparison of enzyme product from three *T. reesei* integrated expression strains. The experimental conditions are described in Example 1. Protein ratios differ across transformants and can be quantified as a percentage of the total integrated peak area. The "EGLs" marks the summed area of endoglucanase peaks. EndoH was added to the protein sample in small amounts as a reagent for HPLC analysis.

Figures 48A-48B: Saccharification increased in xylose monomer yield in response to hemicellulase addition to the enzyme composition produced by an integrated strain at 7 mg total protein per gram glucan + xylan in ammonia pretreated cob. Figure 48A: Constituent *Fusarium verticillioides* hemicellulases in the enzyme composition produced by the integrated strain. Figure 48B: Hemicellulases from other fungi.

Figures 49A-49B: Saccharification increased in glucose monomer yield in response to hemicellulase addition to the enzyme composition produced by an integrated strain at 7 mg total protein per gram glucan + xylan in ammonia pretreated cob. Figure 49A: Constituent *Fusarium verticillioides* hemicellulases in the enzyme composition produced by the integrated strain. Figure 49B: Hemicellulases from other fungi.

Figures 50A-50B: Saccharification increased in arabinose monomer yield in response to hemicellulase addition to the enzyme composition produced by an integrated strain at 7 mg total protein per gram glucan + xylan in ammonia pretreated cob. Figure 50A: Constituent *Fusarium verticillioides* hemicellulases in the enzyme composition produced by an integrated strain. Figure 50B: Hemicellulases from other fungi.

Figure 51: A graphical presentation of the saccharification performance across pretreatment conditions. The X-axis corresponds to the experimental results listed in Table 12. Yields are calculated based on the theoretical amounts of glucan or xylan available in the raw switchgrass. All yields are based on monomeric sugars released after 3 days of saccharification with the enzyme cocktail.

Figure 52: Amino acid sequence alignment of a number of GH39 β -xylosidases. Underlined residues in bold face are the predicted catalytic general acid-base residue (marked with "A" above the alignment) and catalytic nucleophile residue (marked with "N" above the alignment). Underlined residues in normal face in the bottom two sequences are within 4 angstroms of the substrate in the active sites of the respective 3D structures (pdb: 1uhv and 2bs9, respectively). Underlined residues in the Fv39A sequence are predicted to be within 4 angstroms of a bound substrate in the active site.

Figure 53: Amino acid sequence alignment of a number of GH43 family hydrolases. Amino acid residues highly conserved among members of the family are shown underlined and in bold face.

Figure 54: Amino acid sequence alignment of a number of GH51 family enzymes. Amino acid residues highly conserved among members of the family are shown underlined and in bold type.

Figures 55A-55B: Amino acid sequence alignments of a number of GH10 and GH11 family endoxylanases. Figure 55A; Alignment of GH10 family xylanases. Underlined residues in bold face are the the catalytic nucleophile residues (marked with "N" above the alignment). Figure 55B; Alignment of GH11 family xylanases. Underlined residues in bold face are the the catalytic nucleophile residues and general acid base residues (marked with "N" and "A", respectively, above the alignment).

Figures 56A-56B: Saccharification of dilute ammonia pretreated switchgrass with various enzyme blends/compositions; Figure 56A: glucan conversion; Figure 56B: xylan conversion. The numbers below the figures on the x-axis refer to the amount of total mg of each protein in a given blend/composition per g of glucan or xylan, as described in Example 13.

Figures 57A-57B: Saccharification of dilute ammonia pretreated switchgrass with an enzyme composition produced by integrated strain H3A; Figure 57A: glucan conversion; Figure 57B: xylan conversion. Experimental conditions are described in Example 14.

Figures 58A-58C: Saccharification of steam-expanded sugarcane bagasse with an enzyme composition produced by integrated strain H3A at different enzyme doses; Figure 58A: glucan conversion; Figure 58B: xylan conversion; Figure 58C: 3-day glucan and xylan conversions. Experimental conditions are described in Example 17.

Figures 59A-59C: Saccharification of dilute-acid pretreated corn fiber with various enzymes or enzyme blends; Figure 59A: glucan conversion; Figure 59B: xylan conversion; Figure 59C: 5-day glucan and xylan conversions. The adjusted sugar (glucose or xylose) reflected the sugar being produced from the enzymatic step minus the starting sugar levels. Ratios shown along the x-axis represent the enzyme dose in mg of total protein per gram of cellulose. Experimental conditions are described in Example 18.

Figures 60A-60B: Figure 60A: Deduced cDNA for Pa51A (SEQ ID NO:46). Figure 60B: Codon optimized cDNA for Pa51A (SEQ ID NO:47).

Figure 61: Coding sequence for a construct comprising a CBH1 signal sequence (underlined) upstream of genomic DNA encoding mature Gz43A (SEQ ID NO:48).

Figure 62: Coding sequence for a construct comprising a CBH1 signal sequence (underlined) upstream of genomic DNA encoding mature Fo43A (SEQ ID NO:49).

Figure 63: Codon optimized coding sequence for a construct comprising a CBH1 signal sequence (underlined) upstream of codon optimized DNA encoding mature Pf51A (SEQ ID NO:50).

Figure 64: Amino acid sequence alignment of a number of GH3 family hydrolases. Amino acid residues highly conserved among members of the family are shown underlined and in bold face type. The box marks the predicted catalytic residue with flanking residues predicted to be involved in substrate binding.

Figure 65: Amino acid sequence alignment of two representative *Fusarium* GH30 family hydrolases. Amino acid residues that are conserved among members of the family are shown underlined and in bold face type.

6. DETAILED DESCRIPTION

[0042] Enzymes have traditionally been classified by substrate specificity and reaction products. In the pre-genomic era, function was regarded as the most amenable (and perhaps most useful) basis for comparing enzymes and assays for various enzymatic activities have been well-developed for many years, resulting in the familiar EC classification scheme. Cellulases and other glycosyl hydrolases, which act upon glycosidic bonds between two carbohydrate moieties (or a carbohydrate and non-carbohydrate moiety--as occurs in nitrophenol-glycoside derivatives) are, under this classification scheme, designated as EC 3.2.1.-, with the final number indicating the exact type of bond cleaved. For example, according to this scheme an endo-acting cellulase (1,4- β -endoglucanase) is designated EC 3.2.1.4.

[0043] With the advent of widespread genome sequencing projects, sequencing data have facilitated analyses and comparison of related genes and proteins. Additionally, a growing number of enzymes capable of acting on carbohydrate moieties (*i.e.*, carbohydrases) have been crystallized and their 3-D structures solved. Such analyses have identified discreet families of enzymes with related sequence, which contain conserved three-dimensional folds that can be predicted based on their amino acid sequence. Further, it has been shown that enzymes with the same or similar three-dimensional folds exhibit the same or similar stereospecificity of hydrolysis, even when catalyzing different reactions (Henrissat et al., 1998, FEBS Lett 425(2): 352-4; Coutinho and Henrissat, 1999, in Genetics, biochemistry and ecology of cellulose degradation. T. Kimura. Tokyo, Uni Publishers Co: 15-23.).

[0044] These findings form the basis of a sequence-based classification of carbohydrase modules, which is available in the form of an internet database, the Carbohydrate-Active enZYme server (CAZy), available at <http://afmb.cnrs-mrs.fr/CAZY/index.html> (Carbohydrate-active enzymes: an integrated database approach. See Cantarel et al., 2009, Nucleic Acids Res. 37 (Database issue):D233-38).

[0045] CAZy defines four major classes of carbohydrases distinguishable by the type of reaction catalyzed: Glycosyl Hydrolases (GH's), Glycosyltransferases (GT's), Polysaccharide Lyases (PL's), and Carbohydrate Esterases (CE's). The enzymes of the disclosure are glycosyl hydrolases. GH's are a group of enzymes that hydrolyze the glycosidic bond between two or more carbohydrates, or between a carbohydrate and a non-carbohydrate moiety. A classification system for glycosyl hydrolases, grouped by sequence similarity, has led to the definition of over 85 different families. This classification is available on the CAZy web site.

[0046] The enzymes of the disclosure belong, *inter alia*, to the glycosyl hydrolase families 3, 10, 11, 30, 39, 43, and/or 51.

[0047] Glycoside hydrolase family 3 ("GH3") enzymes include, *e.g.*, β -glucosidase (EC:3.2.1.21); β -xylosidase (EC:3.2.1.37); N-acetyl β -glucosaminidase (EC:3.2.1.52); glucan β -1,3-glucosidase (EC:3.2.1.58); cellodextrinase (EC:3.2.1.74); exo-1,3-1,4-glucanase (EC:3.2.1); and β -galactosidase (EC 3.2.1.23). For example, GH3 enzymes can be those that have β -glucosidase, β -xylosidase, N-acetyl β -glucosaminidase, glucan β -1,3-glucosidase,

cellodextrinase, exo-1,3-1,4-glucanase, and/or β -galactosidase activity. Generally, GH3 enzymes are globular proteins and can consist of two or more subdomains. A catalytic residue has been identified as an aspartate residue that, in β -glucosidases, located in the N-terminal third of the peptide and sits within the amino acid fragment SDW (Li et al. 2001, Biochem. J. 355:835-840). The corresponding sequence in Bgl1 from *T. reesei* is T266D267W268 (counting from the methionine at the starting position), with the catalytic residue aspartate being the D267. The hydroxyl/aspartate sequence is also conserved in the GH3 β -xylosidases tested. For example, the corresponding sequence in *T. reesei* Bxl1 is S310D311 and the corresponding sequence in Fv3A is S290D291.

[0048] Glycoside hydrolase family 39 ("GH39") enzymes have α -L-iduronidase (EC:3.2.1.76) or β -xylosidase (EC:3.2.1.37) activity. The three-dimensional structure of two GH39 β -xylosidases, from *Thermoanaerobacterium saccharolyticum* (Uniprot Accession No. P36906) and *Geobacillus stearothermophilus* (Uniprot Accession No. Q9ZFM2), have been solved (see Yang et al. J. Mol. Biol. 2004, 335(1):155-65 and Czjzek et al., J. Mol. Biol. 2005, 353(4):838-46). The most highly conserved regions in these enzymes are located in their N-terminal sections, which have a classic (α/β)₈ TIM barrel fold with the two key active site glutamic acids located at the C-terminal ends of β -strands 4 (acid/base) and 7 (nucleophile). Fv39A residues E168 and E272 are predicted to function as catalytic acid-base and nucleophile, respectively, based on a sequence alignment of the abovementioned GH39 β -xylosidases from *Thermoanaerobacterium saccharolyticum* and *Geobacillus stearothermophilus* with Fv39A.

[0049] Glycoside hydrolase family 43 ("GH43") enzymes include, e.g., L- α -arabinofuranosidase (EC 3.2.1.55); β -xylosidase (EC 3.2.1.37); endo-arabinanase (EC 3.2.1.99); and/or galactan 1,3- β -galactosidase (EC 3.2.1.145). For example, GH43 enzymes can have L- α -arabinofuranosidase activity, β -xylosidase activity, endo-arabinanase activity, and/or galactan 1,3- β -galactosidase activity. GH43 family enzymes display a five-bladed- β -propeller-like structure. The propeller-like structure is based upon a five-fold repeat of blades composed of four-stranded β -sheets. The catalytic general base, an aspartate, the catalytic general acid, a glutamate, and an aspartate that modulates the pKa of the general base were identified through the crystal structure of *Cellvibrio japonicus* CjAbn43A, and confirmed by site-directed mutagenesis (see Nurizzo et al. Nat. Struct. Biol. 2002, 9(9) 665-8). The catalytic residues are arranged in three conserved blocks spread widely through the amino acid sequence (Pons et al. Proteins: Structure, Function and Bioinformatics, 2004, 54:424-432). Among the GH43 family enzymes tested for useful activities in biomass hydrolysis, the predicted catalytic residues are shown as the bold and underlined residues in the sequences of Figure 53. The crystal structure of the *Geobacillus stearothermophilus* xylosidase (Brux et al. J. Mol. Bio., 2006, 359:97-109) suggests several additional residues that may be important for substrate binding in this enzyme. Because the GH43 family enzymes tested for biomass hydrolysis had differing substrate preferences, these residues are not fully conserved in the sequences aligned in Figure 53. However among the xylosidases tested, several conserved residues that contribute to substrate binding, either through hydrophobic interaction or through hydrogen bonding, are conserved and are noted by single underlines in Figure 53.

[0050] Glycoside hydrolase family 51 ("GH51") enzymes have L- α -arabinofuranosidase (EC 3.2.1.55) and/or endoglucanase (EC 3.2.1.4) activity. High-resolution crystal structure of a GH51 L- α -arabinofuranosidase from *Geobacillus stearothermophilus* T-6 shows that the enzyme is a hexamer, with each monomer organized into two domains: an 8-barrel (β/α) and a 12-stranded β sandwich with jelly-roll topology (see Hövel et al. EMBO J. 2003, 22(19):4922-4932). It can be expected that the catalytic residues will be acidic and conserved across enzyme sequences in the family. When the amino acid sequences of Fv51A, Pf51A, and Pa51A are aligned with GH51 enzymes of more diverse sequence, 8 acidic residues remain conserved. Those are shown bold and underlined in Figure 54.

[0051] Glycoside hydrolase family 10 ("GH10") enzymes also have an 8-barrel (β/α) structure. They hydrolyze in an endo fashion with a retaining mechanism that uses at least one acidic catalytic residue in a generally acid/base catalysis process (Pell et al., J. Biol. Chem., 2004, 279(10): 9597-9605). Crystal structures of the GH10 xylanases of *Penicillium simplicissimum* (Uniprot P56588) and *Thermoascus aurantiacus* (Uniprot P23360) complexed with substrates in the active sites have been solved (see Schmidt et al. Biochem., 1999, 38:2403-2412; and Lo Leggio et al. FEBS Lett. 2001, 509: 303-308). *Trichoderma reesei* Xyn3 residues that are important for substrate binding and catalysis can be derived from an alignment with the sequences of abovementioned GH10 xylanases from *Penicillium simplicissimum* and *Thermoascus aurantiacus* (Figure 55A). *Trichoderma reesei* Xyn3 residue E282 is predicted to be the catalytic nucleophilic residue, whereas residues E91, N92, K95, Q97, S98, H128, W132, Q135, N175, E176, Y219, Q252, H254, W312, and/or W320 are predicted to be involved in substrate binding and/or catalysis.

[0052] Glycoside hydrolase family 11 ("GH11") enzymes have a β -jelly roll structure. They hydrolyze in an endo fashion with a retaining mechanism that uses at least one acidic catalytic residue in a generally acid/base catalysis process. Several other residues spread throughout their structure may contribute to stabilizing the xylose units in the substrate neighboring the pair of xylose monomers that are cleaved by hydrolysis. Three GH11 family endoxylanases were tested and their sequences are aligned in Figure 55B. E118 (or E86 in mature *T. reesei* Xyn2) and E209 (or E177 in mature *T. reesei* Xyn2) have been identified as catalytic nucleophile and general/acid base residues in *Trichoderma reesei* Xyn2, respectively (see Havukainen et al. Biochem., 1996, 35:9617-24).

[0053] Glycoside hydrolase family 30 ("GH30") enzymes are retaining enzymes having glucosylceramidase (EC 3.2.1.45); β -1,6-glucanase (EC 3.2.1.75); β -xylosidase (EC 3.2.1.37); β -glucosidase (3.2.1.21) activity. The first GH30 crystal structure was the Gaucher disease-related human β -glucocerebrosidase solved by Grabowski, Gatt and Horowitz (Crit Rev Biochem Mol Biol 1990; 25(6) 385-414). GH30 have an (α/β)₈ TIM barrel fold with the two key active site glutamic acids located at the C-terminal ends of β -strands 4 (acid/base) and 7 (nucleophile) (Henrissat B, et al. Proc Natl Acad Sci U S A, 92(15):7090-4, 1995; Jordan et al., Applied Microbiol Biotechnol, 86:1647, 2010). Glutamate 162 of Fv30A is conserved in 14 of 14 aligned GH30 proteins (13 bacterial proteins and one endo- β -xylanase from the fungi *Biospora* accession no. ADG62369) and glutamate 250 of Fv30A is conserved in 10 of the same 14, is

an aspartate in another three and non-acidic in one. There are other moderately conserved acidic residues but no others are as widely conserved.

6.1 Polypeptides of the Disclosure

[0054] The disclosure provides isolated, synthetic or recombinant polypeptides comprising an amino acid sequence having at least about 80%, *e.g.*, at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete (100%) sequence identity to a polypeptide of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, 44, or 45, over a region of at least about 10, *e.g.*, at least about 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, or 350 residues, or over the full length of the immature polypeptide, the full length mature polypeptide, the full length of the conserved domain, and/or the full length CBM. The conserved domain can be a predicted catalytic domain ("CD"). Exemplary polypeptides also include fragments of at least about 10, *e.g.*, at least about 15, 20, 25, 30, 35, 40, 45, 50, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, or 600 residues in length. The fragments can comprise a conserved domain and/or a CBM. Where a fragment comprises a conserved domain and a CBM of an enzyme, the fragment optionally includes a linker separating the two. The linker can be a native linker or a heterologous linker. It is contemplated that the polypeptides of the disclosure can be encoded by a nucleic acid sequence having at least about 85%, about 86%, about 87%, about 88%, about 89%, or about 90% sequence identity to SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, or 41, or by a nucleic acid sequence capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, or 41, or to a fragment thereof. Exemplary nucleic acids of the disclosure are described in Section 6.2 below.

[0055] The polypeptides of the disclosure include proteins having an amino acid sequence with at least 85%, *e.g.*, at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to at least 50 contiguous amino acid residues of the glycosyl hydrolase sequences of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, 44, or 45. For example, a polypeptide of the disclosure can include amino acid sequences having at least 85%, *e.g.*, at least 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to at least 10, *e.g.*, at least 11, 12, 13, 14, 15, 20, 25, 30, 40, 50, 75, 100, 125, 150, 175, 200, 250, 300, or 350 contiguous amino acid residues of the glycosyl hydrolase sequences of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, 44, or 45. The contiguous amino acid sequence corresponds to the conserved domain and/or the CBM and/or the signal sequence.

[0056] Any of the amino acid sequences described herein can be produced together or in conjunction with at least 1, *e.g.*, at least 2, 3, 5, 10, or 20 heterologous amino acids flanking each of the C- and/or N-terminal ends of the specified amino acid sequence, and or deletions of at least 1, *e.g.*, at least 2, 3, 5, 10, or 20 amino acids from the C- and/or N-terminal ends of

an enzyme of the disclosure.

[0057] Other variations also are within the scope of this disclosure. For example, one or more amino acid residues can be modified to increase or decrease the pI of an enzyme. The change of pI value can be achieved by removing a glutamate residue or substituting it with another amino acid residue.

[0058] The disclosure specifically provides an Fv3A, a Pf43A, an Fv43E, an Fv39A, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, an Af43A, a Pf51A, an AfuXyn2, an AfuXyn5, a Fv43D, a Pf43B, Fv43B, a Fv51A, a *Trichoderma reesei* Xyn3, a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Bxl1, and/or a *Trichoderma reesei* Bgl1 polypeptide. A combination of one or more of these enzymes is suitably present in the enzyme blend/composition, for example, one that is non-naturally occurring.

[0059] Fv3A: The amino acid sequence of Fv3A (SEQ ID NO:2) is shown in Figures 1B, 38, and 64. SEQ ID NO:2 is the sequence of the immature Fv3A. Fv3A has a predicted signal sequence corresponding to residues 1 to 23 of SEQ ID NO:2 (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:2. The predicted conserved domains are in boldface type in Figure 1B. Fv3A was shown to have β -xylosidase activity, for example, in an enzymatic assay using *p*-nitrophenyl- β -xylopyranoside, xylobiose, mixed linear xylo-oligomers, branched arabinoxylan oligomers from hemicellulose, or dilute ammonia pretreated corncob as substrates. The predicted catalytic residue is D291, while the flanking residues, S290 and C292, are predicted to be involved in substrate binding (Figure 64). E175 and E213 are conserved across other GH3 enzymes and are predicted to have catalytic functions. As used herein, "an Fv3A 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, or 700 contiguous amino acid residues among residues 24 to 766 of SEQ ID NO:2. An Fv3A polypeptide preferably is unaltered as compared to native Fv3A in residues D291, S290, C292, E175, and E213. An Fv3A 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 Fv3A, *Trichoderma reesei* Bxl1 and/or *Trichoderma reesei* Bgl1, as shown in the alignment of Figure 64. An Fv3A polypeptide suitably comprises the entire predicted conserved domain of native Fv3A as shown in Figure 1B. An exemplary Fv3A polypeptide 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 Fv3A sequence as shown in Figure 1B. The Fv3A polypeptide preferably has β -xylosidase activity.

[0060] Accordingly an Fv3A polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the amino acid sequence of SEQ ID NO:2, or to residues (i) 24-766, (ii) 73-321, (iii) 73-394, (iv) 395-622, (v) 24-622, or (vi) 73-622 of SEQ ID NO:2. The polypeptide suitably has β -xylosidase

activity.

[0061] Pf43A: The amino acid sequence of Pf43A (SEQ ID NO:4) is shown in Figures 2B and 53. SEQ ID NO:4 is the sequence of the immature Pf43A. Pf43A has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:4 (underlined in Figure 2B); 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:4. The predicted conserved domain is in boldface type, the predicted CBM is in uppercase type, and the predicted linker separating the CD and CBM is in italics in Figure 2B. Pf43A has been shown to have β -xylosidase activity, in, for example, an enzymatic assay using *p*-nitrophenyl- β -xylopyranoside, xylobiose, mixed linear xylo-oligomers, or ammonia pretreated corncob as substrates. The predicted catalytic residues include either D32 or D60, D145, and E206. The C-terminal region underlined in Figure 53 is the predicted CBM. As used herein, "a Pf43A polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 contiguous amino acid residues among residues 21 to 445 of SEQ ID NO:4. A Pf43A polypeptide preferably is unaltered as compared to the native Pf43A in residues D32 or D60, D145, and E206. A Pf43A is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are found conserved across a family of proteins including Pf43A and 1, 2, 3, 4, 5, 6, 7, or all 8 of other amino acid sequences in the alignment of Figure 53. A Pf43A polypeptide suitably comprises two or more or all of the following domains: (1) the predicted CBM, (2) the predicted conserved domain, and (3) the linker of Pf43A as shown in Figure 2B. An exemplary Pf43A polypeptide 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 2B. The Pf43A polypeptide preferably has β -xylosidase activity.

[0062] Accordingly a Pf43A polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the amino acid sequence of SEQ ID NO:4, or to residues (i) 21-445, (ii) 21-301, (iii) 21-323, (iv) 21-444, (v) 302-444, (vi) 302-445, (vii) 324-444, or (viii) 324-445 of SEQ ID NO:4. The polypeptide suitably has β -xylosidase activity.

[0063] Fv43E: The amino acid sequence of Fv43E (SEQ ID NO:6) is shown in Figures 3B and 53. SEQ ID NO:6 is the sequence of the immature Fv43E. Fv43E has a predicted signal sequence corresponding to residues 1 to 18 of SEQ ID NO:6 (underlined in Figure 3B); 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:6. The predicted conserved domain is marked in boldface type in Figure 3B. Fv43E was shown to have β -xylosidase activity, in, for example, enzymatic assay using 4-nitrophenyl- β -D-xylopyranoside, xylobiose, and mixed, linear xylo-oligomers, or ammonia pretreated corncob as substrates. The predicted catalytic residues include either D40 or D71, D155, and E241. As used herein, "an Fv43E polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity

to at least 50, 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:6. An Fv43E polypeptide preferably is unaltered as compared to the native Fv43E in residues D40 or D71, D155, and E241. An Fv43E polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are found to be conserved among a family of enzymes including Fv43E, and 1, 2, 3, 4, 5, 6, 7, or all other 8 amino acid sequences in the alignment of Figure 53. An Fv43E polypeptide suitably comprises the entire predicted conserved domain of native Fv43E as shown in Figure 3B. An exemplary Fv43E polypeptide 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 mature Fv43E sequence as shown in Figure 3B. The Fv43E polypeptide preferably has β -xylosidase activity.

[0064] Accordingly, an Fv43E polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:6, or to residues (i) 19-530, (ii) 29-530, (iii) 19-300, or (iv) 29-300 of SEQ ID NO:6. The polypeptide suitably has β -xylosidase activity.

[0065] Fv39A: The amino acid sequence of Fv39A (SEQ ID NO:8) is shown in Figures 4B and 52. SEQ ID NO:8 is the sequence of the immature Fv39A. Fv39A has a predicted signal sequence corresponding to residues 1 to 19 of SEQ ID NO:8 (underlined in Figure 4B); 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:8. The predicted conserved domain is shown in boldface type in Figure 4B. Fv39A was shown to have β -xylosidase activity in, for example, an enzymatic assay using *p*-nitrophenyl- β -xylopyranoside, xylobiose or mixed, linear xylo-oligomers as substrates. Fv39A residues E168 and E272 are predicted to function as catalytic acid-base and nucleophile, respectively, based on a sequence alignment of the above-mentioned GH39 xylosidases from *Thermoanaerobacterium saccharolyticum* (Uniprot Accession No. P36906) and *Geobacillus stearothermophilus* (Uniprot Accession No. Q9ZFM2) with Fv39A. As used herein, "an Fv39A polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 contiguous amino acid residues among residues 20 to 439 of SEQ ID NO:8. An Fv39A polypeptide preferably is unaltered as compared to native Fv39A in residues E168 and E272. An Fv39A polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among a family of enzymes including Fv39A and xylosidases from *Thermoanaerobacterium saccharolyticum* and *Geobacillus stearothermophilus* (see above). An Fv39A polypeptide suitably comprises the entire predicted conserved domain of native Fv39A as shown in Figure 4B. An exemplary Fv39A polypeptide 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 4B. The Fv39A polypeptide preferably has β -xylosidase activity.

[0066] Accordingly, an Fv39A polypeptide suitably comprises an amino acid sequence with at

least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:8, or to residues (i) 20-439, (ii) 20-291, (iii) 145-291, or (iv) 145-439 of SEQ ID NO:8. The polypeptide suitably has β -xylosidase activity.

[0067] Fv43A: The amino acid sequence of Fv43A (SEQ ID NO:10) is provided in Figures 5B and 53. SEQ ID NO:10 is the sequence of the immature Fv43A. Fv43A has a predicted signal sequence corresponding to residues 1 to 22 of SEQ ID NO:10 (underlined in Figure 5B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 23 to 449 of SEQ ID NO:10. In Figure 5B, the predicted conserved domain is in boldface type, the predicted CBM is in uppercase type, and the predicted linker separating the CD and CBM is in italics. Fv43A was shown to have β -xylosidase activity in, for example, an enzymatic assay using 4-nitrophenyl- β -D-xylopyranoside, xylobiose, mixed, linear xylo-oligomers, branched arabinoxylan oligomers from hemicellulose, and/or linear xylo-oligomers as substrates. The predicted catalytic residues including either D34 or D62, D148, and E209. As used herein, "an Fv43A polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 contiguous amino acid residues among residues 23 to 449 of SEQ ID NO:10. An Fv43A polypeptide preferably is unaltered, as compared to native Fv43A, at residues D34 or D62, D148, and E209. An Fv43A polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among a family of enzymes including Fv43A and 1, 2, 3, 4, 5, 6, 7, 8, or all 9 other amino acid sequences in the alignment of Figure 53. An Fv43A polypeptide suitably comprises the entire predicted CBM of native Fv43A, and/or the entire predicted conserved domain of native Fv43A, and/or the linker of Fv43A as shown in Figure 5B. An exemplary Fv43A polypeptide 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 Fv43A sequence as shown in Figure 5B. The Fv45A polypeptide preferably has β -xylosidase activity.

[0068] Accordingly an Fv43A polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:10, or to residues (i) 23-449, (ii) 23-302, (iii) 23-320, (iv) 23-448, (v) 303-448, (vi) 303-449, (vii) 321-448, or (viii) 321-449 of SEQ ID NO:10. The polypeptide suitably has β -xylosidase activity.

[0069] Fv43B: The amino acid sequence of Fv43B (SEQ ID NO:12) is shown in Figures 6B and 53. 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 6B); 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 conserved domain is in boldface type in Figure 6B. Fv43B was shown to have both β -xylosidase and L- α -arabinofuranosidase activities, in, for example, a first enzymatic assay using 4-nitrophenyl- β -D-xylopyranoside and p-nitrophenyl- α -L-arabinofuranoside as substrates. It was shown, in a second enzymatic assay, to catalyze the release of arabinose from branched arabino-

xylooligomers and to catalyze the increased xylose release from oligomer mixtures in the presence of other xylosidase enzymes. The predicted catalytic residues include either D38 or D68, D151, and E236. As used herein, "an Fv43B polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, or 550 contiguous amino acid residues among residues 17 to 574 of SEQ ID NO:12. An Fv43B polypeptide preferably is unaltered, as compared to native Fv43B, at residues D38 or D68, D151, and E236. An Fv43B polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among a family of enzymes including Fv43B and 1, 2, 3, 4, 5, 6, 7, 8, or all 9 other amino acid sequences in the alignment of Figure 53. An Fv43B polypeptide suitably comprises the entire predicted conserved domain of native Fv43B as shown in Figures 6B and 53. An exemplary Fv43B polypeptide 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 6B. The Fv43B polypeptide preferably has β -xylosidase activity, L- α -arabinofuranosidase activity, or both β -xylosidase and L- α -arabinofuranosidase activities.

[0070] Accordingly, an Fv43B polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:12, or to residues (i) 17-574, (ii) 27-574, (iii) 17-303, or (iv) 27-303 of SEQ ID NO:12. The polypeptide suitably has β -xylosidase activity, L- α -arabinofuranosidase activity, or both β -xylosidase and L- α -arabinofuranosidase activities.

[0071] Pa51A: The amino acid sequence of Pa51A (SEQ ID NO:14) is shown in Figures 7B and 54. SEQ ID NO:14 is the sequence of the immature Pa51A. Pa51A has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:14 (underlined in Figure 7B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 21 to 676 of SEQ ID NO:14. The predicted L- α -arabinofuranosidase conserved domain is in boldface type in Figure 7B. Pa51A was shown to have both β -xylosidase activity and L- α -arabinofuranosidase activity in, for example, enzymatic assays using artificial substrates *p*-nitrophenyl- β -xylopyranoside and *p*-nitrophenyl- α -L-arabinofuranoside. It was shown to catalyze the release of arabinose from branched arabinoxylo oligomers and to catalyze the increased xylose release from oligomer mixtures in the presence of other xylosidase enzymes. Conserved acidic residues include E43, D50, E257, E296, E340, E370, E485, and E493. As used herein, "a Pa51A polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, or 650 contiguous amino acid residues among residues 21 to 676 of SEQ ID NO:14. A Pa51A polypeptide preferably is unaltered, as compared to native Pa51A, at residues E43, D50, E257, E296, E340, E370, E485, and E493. A Pa51A polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among a group of enzymes including Pa51A, Fv51A, and Pf51A, as shown in the alignment of Figure 54.

A Pa51A polypeptide suitably comprises the predicted conserved domain of native Pa51A as shown in Figure 7B. An exemplary Pa51A polypeptide 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 Pa51A sequence as shown in Figure 7B. The Pa51A polypeptide preferably has β -xylosidase activity, L- α -arabinofuranosidase activity, or both β -xylosidase and L- α -arabinofuranosidase activities.

[0072] Accordingly, a Pa51A polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:14, or to residues (i) 21-676, (ii) 21-652, (iii) 469-652, or (iv) 469-676 of SEQ ID NO:14. The polypeptide suitably has β -xylosidase activity, L- α -arabinofuranosidase activity, or both β -xylosidase and L- α -arabinofuranosidase activities.

[0073] Gz43A: The amino acid sequence of Gz43A (SEQ ID NO:16) is shown in Figures 8B and 53. SEQ ID NO:16 is the sequence of the immature Gz43A. Gz43A has a predicted signal sequence corresponding to residues 1 to 18 of SEQ ID NO:16 (underlined in Figure 8B); 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:16. The predicted conserved domain is in boldface type in Figure 8B. Gz43A was shown to have β -xylosidase activity in, for example, an enzymatic assay using *p*-nitrophenyl- β -xylopyranoside, xylobiose or mixed, and/or linear xylo-oligomers as substrates. The predicted catalytic residues include either D33 or D68, D154, and E243. As used herein, "a Gz43A polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, or 300 contiguous amino acid residues among residues 19 to 340 of SEQ ID NO:16. A Gz43A polypeptide preferably is unaltered, as compared to native Gz43A, at residues D33 or D68, D154, and E243. A Gz43A polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among a group of enzymes including Gz43A and 1, 2, 3, 4, 5, 6, 7, 8 or all 9 other amino acid sequences in the alignment of Figure 53. A Gz43A polypeptide suitably comprises the predicted conserved domain of native Gz43A as shown in Figure 8B. An exemplary Gz43A polypeptide 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 Gz43A sequence as shown in Figure 8B. The Gz43A polypeptide preferably has β -xylosidase activity.

[0074] Accordingly a Gz43A polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:16, or to residues (i) 19-340, (ii) 53-340, (iii) 19-383, or (iv) 53-383 of SEQ ID NO:16. The polypeptide suitably has β -xylosidase activity.

[0075] Fo43A: The amino acid sequence of Fo43A (SEQ ID NO:18) is shown in Figures 9B and 53. SEQ ID NO:18 is the sequence of the immature Fo43A. Fo43A has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:18 (underlined in Figure 9B); 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:18. The predicted conserved domain is in boldface type in Figure 9B. Fo43A was shown to have β -xylosidase activity in, for example, an enzymatic assay using *p*-nitrophenyl- β -xylopyranoside, xylobiose and/or mixed, linear xylo-oligomers as substrates. The predicted catalytic residues include either D37 or D72, D159, and E251. As used herein, "an Fo43A polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, or 300 contiguous amino acid residues among residues 18 to 344 of SEQ ID NO:18. An Fo43A polypeptide preferably is unaltered, as compared to native Fo43A, at residues D37 or D72, D159, and E251. An Fo43A polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among a group of enzymes including Fo43A and 1, 2, 3, 4, 5, 6, 7, 8 or all 9 other amino acid sequences in the alignment of Figure 53. An Fo43A polypeptide suitably comprises the predicted conserved domain of native Fo43A as shown in Figure 9B. An exemplary Fo43A polypeptide 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 9B. The Fo43A polypeptide preferably has β -xylosidase activity.

[0076] Accordingly an Fo43A polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:18, or to residues (i) 21-341, (ii) 107-341, (iii) 21-348, or (iv) 107-348 of SEQ ID NO:18. The polypeptide suitably has β -xylosidase activity.

[0077] Af43A: The amino acid sequence of Af43A (SEQ ID NO:20) is shown in Figures 10B and 53. SEQ ID NO:20 is the sequence of the immature Af43A. The predicted conserved domain is in boldface type in Figure 10B. Af43A was shown to have L- α -arabinofuranosidase activity in, for example, an enzymatic assay using *p*-nitrophenyl- α -L-arabinofuranoside as a substrate. Af43A was shown to catalyze the release of arabinose from the set of oligomers released from hemicellulose via the action of endoxylanase. The predicted catalytic residues include either D26 or D58, D139, and E227. As used herein, "an Af43A polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, or 300 contiguous amino acid residues of SEQ ID NO:20. An Af43A polypeptide preferably is unaltered, as compared to native Af43A, at residues D26 or D58, D139, and E227. An Af43A polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among a group of enzymes including Af43A and 1, 2, 3, 4, 5, 6, 7, 8, or all 9 other amino acid sequences in the alignment of Figure 53. An Af43A polypeptide suitably comprises the predicted conserved domain of native Af43A as shown in Figure 10B. An exemplary Af43A polypeptide comprises a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:20. The Af43A polypeptide preferably has L- α -arabinofuranosidase activity.

[0078] Accordingly an Af43A polypeptide suitably comprises an amino acid sequence with at

least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:20, or to residues (i)15-558, or (ii)15-295 of SEQ ID NO:20. The polypeptide suitably has L- α -arabinofuranosidase activity.

[0079] Pf51A: The amino acid sequence of Pf51A (SEQ ID NO:22) is shown in Figures 11B and 54. SEQ ID NO:22 is the sequence of the immature Pf51A. Pf51A has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:22 (underlined in Figure 11B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 21 to 642 of SEQ ID NO:22. The predicted L- α -arabinofuranosidase conserved domain is in boldface type in Figure 11B. Pf51A was shown to have L- α -arabinofuranosidase activity in, for example, an enzymatic assay using 4-nitrophenyl- α -L-arabinofuranoside as a substrate. Pf51A was shown to catalyze the release of arabinose from the set of oligomers released from hemicellulose via the action of endoxylanase. The predicted conserved acidic residues include E43, D50, E248, E287, E331, E360, E472, and E480. As used herein, "a Pf51A polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, or 600 contiguous amino acid residues among residues 21 to 642 of SEQ ID NO:22. A Pf51A polypeptide preferably is unaltered, as compared to native Pf51A, at residues E43, D50, E248, E287, E331, E360, E472, and E480. A Pf51A polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among Pf51A, Pa51A, and Fv51A, as shown in the alignment of Figure 54. A Pf51A polypeptide suitably comprises the predicted conserved domain of native Pf51A shown in Figure 11B. An exemplary Pf51A polypeptide 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 Pf51A sequence shown in Figure 11B. The Pf51A polypeptide preferably has L- α -arabinofuranosidase activity.

[0080] Accordingly a Pf51A polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:22, or to residues (i) 21-632, (ii) 461-632, (iii) 21-642, or (iv) 461-642 of SEQ ID NO:22. The polypeptide has L- α -arabinofuranosidase activity.

[0081] AfuXyn2: The amino acid sequence of AfuXyn2 (SEQ ID NO:24) is shown in Figures 12B and 55B. SEQ ID NO:24 is the sequence of the immature AfuXyn2. AfuXyn2 has a predicted signal sequence corresponding to residues 1 to 18 of SEQ ID NO:24 (underlined in Figure 12B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 19 to 228 of SEQ ID NO:24. The predicted GH11 conserved domain is in boldface type in Figure 12B. AfuXyn2 was shown to have endoxylanase activity indirectly by observing its ability to catalyze the increased xylose monomer production in the presence of xylobiosidase when the enzymes act on pretreated biomass or on isolated hemicellulose. The conserved catalytic residues include E124, E129, and E215. As used herein, "an AfuXyn2 polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%,

93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, or 200 contiguous amino acid residues among residues 19 to 228 of SEQ ID NO:24. An AfuXyn2 polypeptide preferably is unaltered, as compared to native AfuXyn2, at residues E124, E129 and E215. An AfuXyn2 polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among AfuXyn2, AfuXyn5, and *Trichoderma reesei* Xyn2, as shown in the alignment of Figure 55B. An AfuXyn2 polypeptide suitably comprises the entire predicted conserved domain of native AfuXyn2 shown in Figure 12B. An exemplary AfuXyn2 polypeptide 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 AfuXyn2 sequence shown in Figure 12B. The AfuXyn2 polypeptide preferably has xylanase activity.

[0082] AfuXyn5: The amino acid sequence of AfuXyn5 (SEQ ID NO:26) is shown in Figures 13B and 55B. SEQ ID NO:26 is the sequence of the immature AfuXyn5. AfuXyn5 has a predicted signal sequence corresponding to residues 1 to 19 of SEQ ID NO:26 (underlined in Figure 13B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 20 to 313 of SEQ ID NO:26. The predicted GH11 conserved domains are in boldface type in Figure 13B. AfuXyn5 was shown to have endoxylanase activity indirectly by observing its ability to catalyze increased xylose monomer production in the presence of xylobiosidase when the enzymes act on pretreated biomass or on isolated hemicellulose. The conserved catalytic residues include E119, E124, and E210. The predicted CBM is near the C-terminal end, characterized by numerous hydrophobic residues and follows the long serine-, threonine-rich series of amino acids. The region is shown underlined in Figure 55B. As used herein, "an AfuXyn5 polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, or 275 contiguous amino acid residues among residues 20 to 313 of SEQ ID NO:26. An AfuXyn5 polypeptide preferably is unaltered, as compared to native AfuXyn5, at residues E119, E120, and E210. An AfuXyn5 polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among AfuXyn5, AfuXyn2, and *Trichoderma reesei* Xyn2, as shown in the alignment of Figure 55B. An AfuXyn5 polypeptide suitably comprises the entire predicted CBM of native AfuXyn5 and/or the entire predicted conserved domain of native AfuXyn5 (underlined) shown in Figure 13B. An exemplary AfuXyn5 polypeptide 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 AfuXyn5 sequence shown in Figure 13B. The AfuXyn5 polypeptide preferably has xylanase activity.

[0083] Fv43D: The amino acid sequence of Fv43D (SEQ ID NO:28) is shown in Figures 14B and 53. SEQ ID NO:28 is the sequence of the immature Fv43D. Fv43D has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:28 (underlined in Figure 14B); 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:28. The predicted conserved domain is in boldface type in Figure 14B. Fv43D was shown to have β -xylosidase activity in, for example, an

enzymatic assay using p-nitrophenyl- β -xylopyranoside, xylobiose, and/or mixed, linear xylo-oligomers as substrates. The predicted catalytic residues include either D37 or D72, D159, and E251. As used herein, "an Fv43D polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, 300, or 320 contiguous amino acid residues among residues 21 to 350 of SEQ ID NO:28. An Fv43D polypeptide preferably is unaltered, as compared to native Fv43D, at residues D37 or D72, D159, and E251. An Fv43D polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among a group of enzymes including Fv43D and 1, 2, 3, 4, 5, 6, 7, 8, or all 9 other amino acid sequences in the alignment of Figure 53. An Fv43D polypeptide suitably comprises the entire predicted CD of native Fv43D shown in Figure 14B. An exemplary Fv43D polypeptide 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 shown in Figure 14B. The Fv43D polypeptide preferably has β -xylosidase activity.

[0084] Accordingly an Fv43D polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:28, or to residues (i) 20-341, (ii) 21-350, (iii) 107-341, or (iv) 107-350 of SEQ ID NO:28. The polypeptide suitably has β -xylosidase activity.

[0085] Pf43B: The amino acid sequence of Pf43B (SEQ ID NO:30) is shown in Figures 15B and 53. SEQ ID NO:30 is the sequence of the immature Pf43B. Pf43B has a predicted signal sequence corresponding to residues 1 to 20 of SEQ ID NO:30 (underlined in Figure 15B); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 21 to 321 of SEQ ID NO:30. The predicted conserved domain is in boldface type in Figure 15B. Conserved acidic residues within the conserved domain include D32, D61, D148, and E212. Pf43B was shown to have β -xylosidase activity in, for example, an enzymatic assay using p-nitrophenyl- β -xylopyranoside, xylobiose, and/or mixed, linear xylo-oligomers as substrates. As used herein, "a Pf43B polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, or 280 contiguous amino acid residues among residues 21 to 321 of SEQ ID NO:30. A Pf43B polypeptide preferably is unaltered, as compared to native Pf43B, at residues D32, D61, D148, and E212. A Pf43B polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among a group of enzymes including Pf43B and 1, 2, 3, 4, 5, 6, 7, 8, or all 9 other amino acid sequences in the alignment of Figure 53. A Pf43B polypeptide suitably comprises the predicted conserved domain of native Pf43B shown in Figure 15B. An exemplary Pf43B polypeptide 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 Pf43B sequence shown in Figure 15B. The Pf43B polypeptide preferably has β -xylosidase activity.

[0086] Accordingly a Pf43B polypeptide suitably comprises an amino acid sequence with at

least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:30. The polypeptide suitably has β -xylosidase activity.

[0087] Fv51A: The amino acid sequence of Fv51A (SEQ ID NO:32) is shown in Figures 16B and 54. SEQ ID NO:32 is the sequence of the immature Fv51A. Fv51A has a predicted signal sequence corresponding to residues 1 to 19 of SEQ ID NO:32 (underlined in Figure 16B); 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:32. The predicted L- α -arabinofuranosidase conserved domain is in boldface type in Figure 16B. Fv51A was shown to have L- α -arabinofuranosidase activity in, for example, an enzymatic assay using 4-nitrophenyl- α -L-arabinofuranoside as a substrate. Fv51A was shown to catalyze the release of arabinose from the set of oligomers released from hemicellulose via the action of endoxylanase. Conserved residues include E42, D49, E247, E286, E330, E359, E479, and E487. As used herein, "an Fv51A polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, or 625 contiguous amino acid residues among residues 20 to 660 of SEQ ID NO:32. An Fv51A polypeptide preferably is unaltered, as compared to native Fv51A, at residues E42, D49, E247, E286, E330, E359, E479, and E487. An Fv51A polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among Fv51A, Pa51A, and Pf51A, as shown in the alignment of Figure 54. An Fv51A polypeptide suitably comprises the predicted conserved domain of native Fv51A shown in Figure 16B. An exemplary Fv51A polypeptide 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 Fv51A sequence shown in Figure 16B. The Fv51A polypeptide preferably has L- α -arabinofuranosidase activity.

[0088] Accordingly an Fv51A polypeptide suitably comprise an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO:32, or to residues (i) 21-660, (ii) 21-645, (iii) 450-645, or (iv) 450-660 of SEQ ID NO:32. The polypeptide suitably has L- α -arabinofuranosidase activity.

[0089] Xyn3: The amino acid sequence of *Trichoderma reesei*Xyn3 (SEQ ID NO:42) is shown in Figure 21B. SEQ ID NO:42 is the sequence of the immature *Trichoderma reesei* Xyn3. *Trichoderma reesei* Xyn3 has a predicted signal sequence corresponding to residues 1 to 16 of SEQ ID NO:42 (underlined in Figure 21B); 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:42. The predicted conserved domain is in boldface type in Figure 21B. *Trichoderma reesei* Xyn3 was shown to have endoxylanase activity indirectly by observation of its ability to catalyze increased xylose monomer production in the presence of xylobiosidase when the enzymes act on pretreated biomass or on isolated hemicellulose. The conserved catalytic residues include E91, E176, E180, E195, and E282, as determined by alignment with another GH10 family enzyme, the Xys1 delta from *Streptomyces halstedii* (Canals et al., 2003, Act Crystallogr. D Biol. 59:1447-53), which has 33% sequence identity to *Trichoderma reesei* Xyn3. As used

herein, "a *Trichoderma reesei* Xyn3 polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, or 300 contiguous amino acid residues among residues 17 to 347 of SEQ ID NO:42. A *Trichoderma reesei* Xyn3 polypeptide preferably is unaltered, as compared to native *Trichoderma reesei* Xyn3, at residues E91, E176, E180, E195, and E282. A *Trichoderma reesei* Xyn3 polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved between *Trichoderma reesei* Xyn3 and Xys1 delta. A *Trichoderma reesei* Xyn3 polypeptide suitably comprises the entire predicted conserved domain of native *Trichoderma reesei* Xyn3 shown in Figure 21B. An exemplary *Trichoderma reesei* Xyn3 polypeptide 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 *Trichoderma reesei* Xyn3 sequence shown in Figure 21B. The *Trichoderma reesei* Xyn3 polypeptide preferably has xylanase activity.

[0090] Xyn2: The amino acid sequence of *Trichoderma reesei* Xyn2 (SEQ ID NO:43) is shown in Figures 22 and 55B. SEQ ID NO:43 is the sequence of the immature *Trichoderma reesei* Xyn2. *Trichoderma reesei* Xyn2 has a predicted prepropeptide sequence corresponding to residues 1 to 33 of SEQ ID NO:43 (underlined in Figure 22); cleavage of the predicted signal sequence between positions 16 and 17 is predicted to yield a propeptide, which is processed by a kexin-like protease between positions 32 and 33, generating the mature protein having a sequence corresponding to residues 33 to 222 of SEQ ID NO:43. The predicted conserved domain is in boldface type in Figure 22. *Trichoderma reesei* Xyn2 was shown to have endoxylanase activity indirectly by observation of its ability to catalyze an increased xylose monomer production in the presence of xylobiosidase when the enzymes act on pretreated biomass or on isolated hemicellulose. The conserved acidic residues include E118, E123, and E209. As used herein, "a *Trichoderma reesei* Xyn2 polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, or 175 contiguous amino acid residues among residues 33 to 222 of SEQ ID NO:43. A *Trichoderma reesei* Xyn2 polypeptide preferably is unaltered, as compared to a native *Trichoderma reesei* Xyn2, at residues E118, E123, and E209. A *Trichoderma reesei* Xyn2 polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among *Trichoderma reesei* Xyn2, AfuXyn2, and AfuXyn5, as shown in the alignment of Figure 55B. A *Trichoderma reesei* Xyn2 polypeptide suitably comprises the entire predicted conserved domain of native *Trichoderma reesei* Xyn2 shown in Figure 22. An exemplary *Trichoderma reesei* Xyn2 polypeptide 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 *Trichoderma reesei* Xyn2 sequence shown in Figure 22. The *Trichoderma reesei* Xyn2 polypeptide preferably has xylanase activity.

[0091] Bxl1: The amino acid sequence of *Trichoderma reesei* Bxl1 (SEQ ID NO:44) is shown in Figures 23 and 64. SEQ ID NO:44 is the sequence of the immature *Trichoderma reesei* Bxl1. *Trichoderma reesei* Bxl1 has a predicted signal sequence corresponding to residues 1 to 18 of

SEQ ID NO:44 (underlined in Figure 23); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 19 to 797 of SEQ ID NO:44. The predicted conserved domains are in boldface type in Figure 23. *Trichoderma reesei* Bxl1 was shown to have β -xylosidase activity in, for example, an enzymatic assay using *p*-nitrophenyl- β -xylopyranoside, xylobiose and/or mixed, linear xylo-oligomers as substrates. The conserved acidic residues include E193, E234, and D310. As used herein, "a *Trichoderma reesei* Bxl1 polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, or 750 contiguous amino acid residues among residues 17 to 797 of SEQ ID NO:44. A *Trichoderma reesei* Bxl1 polypeptide preferably is unaltered, as compared to a native *Trichoderma reesei* Bxl1, at residues E193, E234, and D310. A *Trichoderma reesei* Bxl1 polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among *Trichoderma reesei* Bxl1, Fv3A, and *Trichoderma reesei* Bgl1, as shown in the alignment of Figure 64. A *Trichoderma reesei* Bxl1 polypeptide suitably comprises the entire predicted conserved domains of native *Trichoderma reesei* Bxl1 shown in Figure 23. An exemplary *Trichoderma reesei* Bxl1 polypeptide 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 *Trichoderma reesei* Bxl1 sequence shown in Figure 23. The *Trichoderma reesei* Bxl1 polypeptide preferably has β -xylosidase activity.

[0092] Accordingly a *Trichoderma reesei* Bxl1 polypeptide suitably comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the amino acid sequence of SEQ ID NO: 44. The polypeptide suitably has β -xylosidase activity.

[0093] Bgl1: The amino acid sequence of *Trichoderma reesei* Bgl1 (SEQ ID NO:45) is shown in Figures 24 and 64. *Trichoderma reesei* Bgl1 has a predicted signal sequence corresponding to residues 1 to 19 of SEQ ID NO:45 (underlined in Figure 24); cleavage of the signal sequence is predicted to yield a mature protein having a sequence corresponding to residues 20 to 744 of SEQ ID NO:45. The predicted conserved domain is in boldface type in Figure 24. *Trichoderma reesei* Bgl1 has been shown to have β -glucosidase activity by observation of a capacity to catalyze the hydrolysis of para-nitrophenyl- β -D-glucopyranoside to produce para-nitrophenol, and a capacity to catalyze the hydrolysis of cellobiose. The conserved acidic residues include D164, E197, and D267. As used herein, "a *Trichoderma reesei* Bgl1 polypeptide" refers to a polypeptide and/or a variant thereof comprising a sequence having at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to at least 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, or 780 contiguous amino acid residues among residues 20 to 744 of SEQ ID NO:45. A *Trichoderma reesei* Bgl1 polypeptide preferably is unaltered, as compared to a native Bgl1, at residues D164, E197, and D267. A *Trichoderma reesei* Bgl1 polypeptide is preferably unaltered in at least 70%, 80%, 90%, 95%, 98%, or 99% of the amino acid residues that are conserved among *Trichoderma reesei* Bgl1, Fv3A, and *Trichoderma*

reesei Bxl1, as shown in the alignment of Figure 64. A *Trichoderma reesei* Bgl1 polypeptide suitably comprises the entire predicted conserved domain of native *Trichoderma reesei* Bgl1 shown in Figure 24. An exemplary *Trichoderma reesei* Bgl1 polypeptide 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 *Trichoderma reesei* Bgl1 sequence shown in Figure 24. The *Trichoderma reesei* Bgl1 polypeptide preferably has β -glucosidase activity.

[0094] Accordingly, the present disclosure provides a number of isolated, synthetic, or recombinant hemicellulolytic polypeptides or variants as described below:

1. (1) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 24 to 766 of SEQ ID NO:2; (ii) 73 to 321 of SEQ ID NO:2; (iii) 73 to 394 of SEQ ID NO:2; (iv) 395 to 622 of SEQ ID NO:2; (v) 24 to 622 of SEQ ID NO:2; or (iv) 73 to 622 of SEQ ID NO:2; the polypeptide preferably has β -xylosidase activity; or
2. (2) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 21 to 445 of SEQ ID NO:4; (ii) 21 to 301 of SEQ ID NO:4; (iii) 21 to 323 of SEQ ID NO:4; (iv) 21 to 444 of SEQ ID NO:4; (v) 302 to 444 of SEQ ID NO:4; (vi) 302 to 445 of SEQ ID NO:4; (vii) 324 to 444 of SEQ ID NO:4; or (viii) 324 to 445 of SEQ ID NO:4; the polypeptide preferably has β -xylosidase activity; or
3. (3) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 19 to 530 of SEQ ID NO:6; (ii) 29 to 530 of SEQ ID NO:6; (iii) 19 to 300 of SEQ ID NO:6; or (iv) 29 to 300 of SEQ ID NO:6; the polypeptide preferably has β -xylosidase activity; or
4. (4) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 20 to 439 of SEQ ID NO:8; (ii) 20 to 291 of SEQ ID NO:8; (iii) 145 to 291 of SEQ ID NO:8; or (iv) 145 to 439 of SEQ ID NO:8; the polypeptide preferably has β -xylosidase activity; or
5. (5) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 23 to 449 of SEQ ID NO:10; (ii) 23 to 302 of SEQ ID NO:10; (iii) 23 to 320 of SEQ ID NO:10; (iv) 23 to 448 of SEQ ID NO:10; (v) 303 to 448 of SEQ ID NO:10; (vi) 303 to 449 of SEQ ID NO:10; (vii) 321 to 448 of SEQ ID NO:10; or (viii) 321 to 449 of SEQ ID NO:10; the polypeptide preferably has β -xylosidase activity; or
6. (6) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 17 to 574 of SEQ ID NO:12; (ii) 27 to 574 of SEQ ID NO:12; (iii) 17 to 303 of SEQ ID NO:12; or (iv) 27 to 303 of SEQ ID NO:12; the polypeptide preferably has both β -xylosidase activity and L- α -arabinofuranosidase activity; or

7. (7) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 21 to 676 of SEQ ID NO:14; (ii) 21 to 652 of SEQ ID NO:14; (iii) 469 to 652 of SEQ ID NO:14; or (iv) 469 to 676 of SEQ ID NO:14; the polypeptide preferably has both β -xylosidase activity and L- α -arabinofuranosidase activity; or
8. (8) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 19 to 340 of SEQ ID NO:16; (ii) 53 to 340 of SEQ ID NO:16; (iii) 19 to 383 of SEQ ID NO:16; or (iv) 53 to 383 of SEQ ID NO:16; the polypeptide preferably has β -xylosidase activity; or
9. (9) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 21 to 341 of SEQ ID NO:18; (ii) 107 to 341 of SEQ ID NO:18; (iii) 21 to 348 of SEQ ID NO:18; or (iv) 107 to 348 of SEQ ID NO:18; the polypeptide preferably has β -xylosidase activity; or
10. (10) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 15 to 558 of SEQ ID NO:20; or (ii) 15 to 295 of SEQ ID NO:20; the polypeptide preferably has L- α -arabinofuranosidase activity; or
11. (11) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 21 to 632 of SEQ ID NO:22; (ii) 461 to 632 of SEQ ID NO:22; (iii) 21 to 642 of SEQ ID NO:22; or (iv) 461 to 642 of SEQ ID NO:22; the polypeptide preferably has L- α -arabinofuranosidase activity; or
12. (12) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 20 to 341 of SEQ ID NO:28; (ii) 21 to 350 of SEQ ID NO:28; (iii) 107 to 341 of SEQ ID NO:28; or (iv) 107 to 350 of SEQ ID NO:28; the polypeptide has β -xylosidase activity; or
13. (13) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 21 to 660 of SEQ ID NO:32; (ii) 21 to 645 of SEQ ID NO:32; (iii) 450 to 645 of SEQ ID NO:32; or (iv) 450 to 660 of SEQ ID NO:32; the polypeptide preferably has L- α -arabinofuranosidase activity.

[0095] The present disclosure provides also compositions (e.g., cellulase compositions, or enzyme blends/compositions) or fermentation broths enriched with one or more of the above-described polypeptides. The enzyme blend/composition is thus a non-naturally-occurring composition. The cellulase composition can be, for example, a filamentous fungal cellulase composition, such as a *Trichoderma* cellulase composition. The fermentation broth can be a fermentation broth of a filamentous fungus, for example, a *Trichoderma*, *Humicola*, *Fusarium*,

Aspergillus, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, or *Chrysosporium* fermentation broth. In particular, the fermentation broth can be, for example, one of *Trichoderma* spp. such as a *Trichoderma reesei*, or *Penicillium* spp., such as a *Penicillium funiculosum*. The fermentation broth can also suitably be a cell-free fermentation broth.

[0096] Additionally the instant disclosure provides host cells that are recombinantly engineered to express a polypeptide described above. The host cells can be, for example, filamentous fungal host cells, such as *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *cochliobolus*, *Pyricularia*, or *Chrysosporium* cells. In particular, the host cells can be, for example, a *Trichoderma* spp. cell (such as a *Trichoderma reesei* cell), or a *Penicillium* cell (such as a *Penicillium funiculosum* cell), an *Aspergillus* cell (such as an *Aspergillus oryzae* or *Aspergillus nidulans* cell), or a *Fusarium* cell (such as a *Fusarium verticilloides* or *Fusarium oxysporum* cell).

6.1.1 Fusion proteins

[0097] The present disclosure also provides a fusion protein that includes a domain of a protein of the present disclosure attached to one or more fusion segments, which are typically heterologous to the protein (*i.e.*, derived from a different source than the protein of the disclosure). Suitable fusion segments include, without limitation, segments that can enhance a protein's stability, provide other desirable biological activity, and/or facilitate purification of the protein (*e.g.*, by affinity chromatography). A suitable fusion segment can be a domain of any size that has the desired function (*e.g.*, imparts increased stability, solubility, action or biological activity; and/or simplifies purification of a protein). Fusion segments can be joined to amino and/or carboxyl termini of the domain(s) of a protein of the present disclosure. The fusion segments can be susceptible to cleavage. There may be some advantage in having this susceptibility, for example, it may enable straight-forward recovery of the protein of interest. Fusion proteins are preferably produced by culturing a recombinant cell transfected with a fusion nucleic acid that encodes a protein, which includes a fusion segment attached to either the carboxyl or amino terminal end, or fusion segments attached to both the carboxyl and amino terminal ends, of a protein, or a domain thereof.

[0098] Accordingly, proteins of the present disclosure also include expression products of gene fusions (*e.g.*, an overexpressed, soluble, and active form of a recombinant protein), of mutagenized genes (*e.g.*, genes having codon modifications to enhance gene transcription and translation), and of truncated genes (*e.g.*, genes having signal sequences removed or substituted with a heterologous signal sequence).

[0099] Glycosyl hydrolases that utilize insoluble substrates are often modular enzymes. They usually comprise catalytic modules appended to one or more non-catalytic carbohydrate-binding domains (CBMs). In nature, CBMs are thought to promote the glycosyl hydrolase's interaction with its target substrate polysaccharide. Thus, the disclosure provides chimeric

enzymes having altered substrate specificity; including, for example, chimeric enzymes having multiple substrates as a result of "spliced-in" heterologous CBMs. The heterologous CBMs of the chimeric enzymes of the disclosure can also be designed to be modular, such that they are appended to a catalytic module or catalytic domain (a "CD", e.g., at an active site), which can likewise be heterologous or homologous to the glycosyl hydrolase.

[0100] Thus, the disclosure provides peptides and polypeptides consisting of, or comprising, CBM/CD modules, which can be homologously paired or joined to form chimeric (heterologous) CBM/CD pairs. Thus, these chimeric polypeptides/peptides can be used to improve or alter the performance of an enzyme of interest. Accordingly, the disclosure provides chimeric enzymes comprising, e.g., at least one CBM of an enzyme of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, 44, or 45. A polypeptide of the disclosure, for example, includes an amino acid sequence comprising the CD and/or CBM of the glycosyl hydrolase sequence of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, 44, or 45. The polypeptide of the disclosure can thus suitably be a fusion protein comprising functional domains from two or more different proteins (e.g., a CBM from one protein linked to a CD from another protein).

[0101] The polypeptides of the disclosure can suitably be obtained and/or used in "substantially pure" form. For example, a polypeptide of the disclosure constitutes at least about 80 wt.% (e.g., at least about 85 wt.%, 90 wt.%, 91 wt.%, 92 wt.%, 93 wt.%, 94 wt.%, 95 wt.%, 96 wt.%, 97 wt.%, 98 wt.%, or 99 wt.%) of the total protein in a given composition, which also includes other ingredients such as a buffer or solution.

[0102] Also, the polypeptides of the disclosure can suitably be obtained and/or used in recombinant culture broths (e.g., a filamentous fungal culture broth). The recombinant culture broths can be non-naturally occurring; for example, the culture broth can be produced by a recombinant host cell that is engineered to express a heterologous polypeptide of the disclosure, or by a recombinant host cell that is engineered to express an endogenous polypeptide of the disclosure in greater or lesser amounts than the endogenous expression levels (e.g., in an amount that is 1-, 2-, 3-, 4-, 5-, or more- fold greater or less than the endogenous expression levels). Furthermore, the polypeptides of the disclosure can suitably be obtained and/or used as recombinant culture broths produced by "integrated" host cell strains that have been engineered to express a plurality of polypeptides of the disclosure in desired ratios. Exemplary desired ratios are described herein, for example, in Section 6.3.4 below.

6.2 Nucleic Acids and Host Cells

[0103] The present disclosure provides nucleic acids encoding a polypeptide of the disclosure, for example one described in Section 6.1 above.

[0104] The disclosure provides isolated, synthetic or recombinant nucleic acids comprising a

nucleic acid sequence having at least about 70%, e.g., at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%; 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%, or complete (100%) sequence identity to a nucleic acid of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 46, 47, 48, 49, or 50, over a region of at least about 10, e.g., at least about 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, or 2000 nucleotides. The present disclosure also provides nucleic acids encoding at least one polypeptide having a hemicellulolytic activity (e.g., a xylanase, β -xylosidase, and/or L- α - arabinofuranosidase activity).

[0105] Nucleic acids of the disclosure also include isolated, synthetic or recombinant nucleic acids encoding an enzyme having the sequence of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, 44, or 45, and subsequences thereof (e.g., a conserved domain or carbohydrate binding domain ("CBM"), and variants thereof. A nucleic acid of the disclosure can, for example, encode the mature portion of a protein of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 43, 44, or 45.

[0106] The disclosure specifically provides a nucleic acid encoding an Fv3A, a Pf43A, an Fv43E, an Fv39A, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, an Af43A, a Pf51A, an AfuXyn2, an AfuXyn5, a Fv43D, a Pf43B, an Fv43B, an Fv51A, a *Trichoderma reesei* Xyn3, a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Bxl1, or a *Trichoderma reesei* Bgl1 polypeptide.

[0107] For example, the disclosure provides an isolated nucleic acid molecule, wherein the nucleic acid molecule encodes:

1. (1) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 24 to 766 of SEQ ID NO:2; (ii) 73 to 321 of SEQ ID NO:2; (iii) 73 to 394 of SEQ ID NO:2; (iv) 395 to 622 of SEQ ID NO:2; (v) 24 to 622 of SEQ ID NO:2; or (iv) 73 to 622 of SEQ ID NO:2; or
2. (2) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 21 to 445 of SEQ ID NO:4; (ii) 21 to 301 of SEQ ID NO:4; (iii) 21 to 323 of SEQ ID NO:4; (iv) 21 to 444 of SEQ ID NO:4; (v) 302 to 444 of SEQ ID NO:4; (vi) 302 to 445 of SEQ ID NO:4; (vii) 324 to 444 of SEQ ID NO:4; or (viii) 324 to 445 of SEQ ID NO:4; or
3. (3) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 19 to 530 of SEQ ID NO:6; (ii) 29 to 530 of SEQ ID NO:6; (iii) 19 to 300 of SEQ ID NO:6; or (iv) 29 to 300 of SEQ ID NO:6; or
4. (4) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at

- least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 20 to 439 of SEQ ID NO:8; (ii) 20 to 291 of SEQ ID NO:8; (iii) 145 to 291 of SEQ ID NO:8; or (iv) 145 to 439 of SEQ ID NO:8; or
5. (5) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 23 to 449 of SEQ ID NO:10; (ii) 23 to 302 of SEQ ID NO:10; (iii) 23 to 320 of SEQ ID NO:10; (iv) 23 to 448 of SEQ ID NO:10; (v) 303 to 448 of SEQ ID NO:10; (vi) 303 to 449 of SEQ ID NO:10; (vii) 321 to 448 of SEQ ID NO:10; or (viii) 321 to 449 of SEQ ID NO:10; or
 6. (6) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 17 to 574 of SEQ ID NO:12; (ii) 27 to 574 of SEQ ID NO:12; (iii) 17 to 303 of SEQ ID NO:12; or (iv) 27 to 303 of SEQ ID NO:12; or
 7. (7) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 21 to 676 of SEQ ID NO:14; (ii) 21 to 652 of SEQ ID NO:14; (iii) 469 to 652 of SEQ ID NO:14; or (iv) 469 to 676 of SEQ ID NO:14; or
 8. (8) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 19 to 340 of SEQ ID NO:16; (ii) 53 to 340 of SEQ ID NO:16; (iii) 19 to 383 of SEQ ID NO:16; or (iv) 53 to 383 of SEQ ID NO:16; or
 9. (9) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 21 to 341 of SEQ ID NO:18; (ii) 107 to 341 of SEQ ID NO:18; (iii) 21 to 348 of SEQ ID NO:18; or (iv) 107 to 348 of SEQ ID NO:18; or
 10. (10) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 15 to 558 of SEQ ID NO:20; or (ii) 15 to 295 of SEQ ID NO:20; or
 11. (11) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 21 to 632 of SEQ ID NO:22; (ii) 461 to 632 of SEQ ID NO:22; (iii) 21 to 642 of SEQ ID NO:22; or (iv) 461 to 642 of SEQ ID NO:22; or
 12. (12) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 20 to 341 of SEQ ID NO:28; (ii) 21 to 350 of SEQ ID NO:28; (iii) 107 to 341 of SEQ ID NO:28; or (iv) 107 to 350 of SEQ ID NO:28; or
 13. (13) a polypeptide comprising an amino acid sequence with at least 90%, at least 95%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence corresponding to positions (i) 21 to 660 of SEQ ID NO:32; (ii) 21 to 645 of SEQ ID NO:32; (iii) 450 to 645 of SEQ ID NO:32; or (iv) 450 to 660 of SEQ ID NO:32.

[0108] The instant disclosure also provides:

1. (1) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:1, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:1, or to a fragment thereof; or
2. (2) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:3, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:3, or to a fragment thereof; or
3. (3) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:5, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:5, or to a fragment thereof; or
4. (4) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:7, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:7, or to a fragment thereof; or
5. (5) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:9, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:9, or to a fragment thereof; or
6. (6) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:11, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:11, or to a fragment thereof; or
7. (7) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:13, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:13, or to a fragment thereof; or
8. (8) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:15, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:15, or to a fragment thereof; or
9. (9) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:17, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:17, or to a fragment thereof; or
10. (10) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:19, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:19, or to a fragment thereof; or
11. (11) a nucleic acid having at least 90% (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:21, or a nucleic acid that

is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:21, or to a fragment thereof; or

12. (12) a nucleic acid having at least 90% (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:27, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:27, or to a fragment thereof; or
13. (13) a nucleic acid having at least 90% (e.g., at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more) sequence identity to SEQ ID NO:31, or a nucleic acid that is capable of hybridizing under high stringency conditions to a complement of SEQ ID NO:31, or to a fragment thereof.

[0109] The disclosure also provides expression cassettes and/or vectors comprising the above-described nucleic acids.

[0110] Suitably, the nucleic acid encoding an enzyme of the disclosure is operably linked to a promoter. Specifically, where recombinant expression in a filamentous fungal host is desired, the promoter can be a filamentous fungal promoter. The nucleic acids can be, for example, under the control of heterologous promoters. The nucleic acids can also be expressed under the control of constitutive or inducible promoters. Examples of promoters that can be used include, but are not limited to, a cellulase promoter, a xylanase promoter, the 1818 promoter (previously identified as a highly expressed protein by EST mapping *Trichoderma*). For example, the promoter can suitably be a cellobiohydrolase, endoglucanase, or β -glucosidase promoter. A particular suitable promoter can be, for example, a *T. reesei* cellobiohydrolase, endoglucanase, or β -glucosidase promoter. For example, the promoter is a cellobiohydrolase I (*cbh1*) promoter. Non-limiting examples of promoters include a *cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl4*, *egl5*, *pkil*, *gpdI*, *xyn1*, or *xyn2* promoter. Additional non-limiting examples of promoters include a *T. reesei* *cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl4*, *egl5*, *pkil*, *gpdI*, *xyn1*, or *xyn2* promoter.

[0111] The present disclosure provides host cells that are engineered to express one or more enzymes of the disclosure. Suitable host cells include cells of any microorganism (e.g., cells of a bacterium, a protist, an alga, a fungus (e.g., a yeast or filamentous fungus), or other microbe), and are preferably cells of a bacterium, a yeast, or a filamentous fungus.

[0112] Suitable host cells of the bacterial genera include, but are not limited to, cells of *Escherichia*, *Bacillus*, *Lactobacillus*, *Pseudomonas*, and *Streptomyces*. Suitable cells of bacterial species include, but are not limited to, cells of *Escherichia coli*, *Bacillus subtilis*, *Bacillus licheniformis*, *Lactobacillus brevis*, *Pseudomonas aeruginosa*, and *Streptomyces lividans*.

[0113] Suitable host cells of the genera of yeast include, but are not limited to, cells of *Saccharomyces*, *Schizosaccharomyces*, *Candida*, *Hansenula*, *Pichia*, *Kluyveromyces*, and *Phaffia*. Suitable cells of yeast species include, but are not limited to, cells of *Saccharomyces*

cerevisiae, *Schizosaccharomyces pombe*, *Candida albicans*, *Hansenula polymorpha*, *Pichia pastoris*, *P. canadensis*, *Kluyveromyces marxianus*, and *Phaffia rhodozyma*.

[0114] Suitable host cells of filamentous fungi include all filamentous forms of the subdivision *Eumycotina*. Suitable cells of filamentous fungal genera include, but are not limited to, cells of *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Corynascus*, *Chaetomium*, *Cryptococcus*, *Filobasidium*, *Fusarium*, *Gibberella*, *Humicola*, *Hypocrea*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Mucor*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Scytalidium*, *Schizophyllum*, *Sporotrichum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, and *Trichoderma*. Suitable cells can also include cells of various anamorph and teleomorph forms for these filamentous fungal genera.

[0115] Suitable cells of filamentous fungal species include, but are not limited to, cells of *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Neurospora intermedia*, *Penicillium purpurogenum*, *Penicillium canescens*, *Penicillium solitum*, *Penicillium funiculosum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, and *Trichoderma viride*.

[0116] The disclosure further provides a recombinant host cell that is engineered to express one or more, two or more, three or more, four or more, or five or more of an Fv3A, a Pf43A, an Fv43E, an Fv39A, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, an Af43A, a Pf51A, an AfuXyn2, an AfuXyn5, an Fv43D, a Pf43B, and an Fv51A polypeptide. The recombinant host cell is, for example, a recombinant *Trichoderma reesei* host cell. In a particular example, the disclosure provides a recombinant fungus, such as a recombinant *Trichoderma reesei*, that is engineered to express one or more, two or more, three or more, four or more, or five or more of an Fv3A, a Pf43A, an Fv43E, an Fv39A, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, an Af43A, a Pf51A, an AfuXyn2, an AfuXyn5, an Fv43D, a Pf43B, and an Fv51A polypeptide. The disclosure provides a recombinant *Trichoderma reesei* host cell engineered to express 1, 2, 3, 4, 5, or more of an Fv3A, a Pf43A, an Fv43E, an Fv39A, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, an Af43A, a Pf51A, an AfuXyn2, an AfuXyn5, an Fv43D, a Pf43B, and an Fv51A polypeptide.

[0117] The disclosure provides a host cell, for example, a recombinant fungal host cell or a recombinant filamentous fungus, engineered to recombinantly express at least one xylanase, at least one β -xylosidase, and one L- α -arabinofuranosidase. The disclosure also provides a recombinant host cell, e.g., a recombinant fungal host cell or a recombinant filamentous fungus such as a recombinant *Trichoderma reesei*, that is engineered to express 1, 2, 3, 4, 5, or more of an Fv3A, a Pf43A, an Fv43E, an Fv39A, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, an Af43A, a Pf51A, an AfuXyn2, an AfuXyn5, an Fv43D, a Pf43B, and an Fv51A polypeptide, in addition to one or more of *Trichoderma reesei* Xyn2, *Trichoderma reesei* Xyn3, *Trichoderma reesei* Bxl1 and/or *Trichoderma reesei* Bgl1. The recombinant host cell is, for example, a *Trichoderma reesei* host cell. The recombinant fungus is, for example, a recombinant *Trichoderma reesei*. The disclosure provides a *Trichoderma reesei* host cell, or a recombinant *Trichoderma reesei* fungus, that is engineered to recombinantly express 1, 2, 3, 4, 5, or more of an Fv3A, a Pf43A, an Fv43E, an Fv39A, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, an Af43A, a Pf51A, an AfuXyn2, an AfuXyn5, an Fv43D, a Pf43B, and an Fv51A polypeptide, in addition to recombinantly express one or more of *Trichoderma reesei* Xyn2, *Trichoderma reesei* Xyn3, *Trichoderma reesei* Bxl1 and/or *Trichoderma reesei* Bgl1.

[0118] The present disclosure also provides a recombinant host cell e.g., a recombinant fungal host cell or a recombinant organism, e.g., a filamentous fungus, such as a recombinant *Trichoderma reesei*, that is engineered to recombinantly express *Trichoderma reesei* Xyn3, *Trichoderma reesei* Bgl1, Fv3A, Fv43D, and Fv51A polypeptides. For example, the recombinant host cell is suitably a *Trichoderma reesei* host cell. The recombinant fungus is suitably a recombinant *Trichoderma reesei*. The disclosure provides, for example, a *Trichoderma reesei* host cell engineered to recombinantly express *Trichoderma reesei* Xyn3, *Trichoderma reesei* Bgl1, Fv3A, Fv43D, and Fv51A polypeptides.

[0119] Additionally the disclosure provides a recombinant host cell or recombinant fungus that is engineered to express an enzyme blend comprising suitable enzymes in ratios suitable for saccharification. The recombinant host cell is, for example, a fungal host cell. The recombinant fungus is, for example, a recombinant *Trichoderma reesei*. Exemplary enzyme ratios/amounts present in suitable enzyme blends are described in Section 6.3.4 below.

[0120] The disclosure further provides transgenic plants comprising a nucleic acid of the disclosure or an expression cassette of the disclosure. The transgenic plant can be, for example, a cereal plant, a corn plant, a potato plant, a tomato plant, a wheat plant, an oilseed plant, a rapeseed plant, a soybean plant, a rice plant, a barley plant, or a tobacco plant.

6.3 Enzyme Blends for Saccharification

[0121] The present disclosure provides a composition comprising an enzyme blend/composition that is capable of breaking down lignocellulose material. Such a multi-enzyme blend/composition comprises at least one polypeptide of the present disclosure, in

combination with one or more additional polypeptides of the present disclosure, or one or more enzymes from other microorganisms, plants, or organisms. Synergistic enzyme combinations and related methods are contemplated. The disclosure includes methods for identifying the optimum ratios of the enzymes included in the blends/compositions for degrading a particular lignocellulosic material. These methods include, e.g., tests to identify the optimum enzyme blend/composition and ratios for efficient conversion of a given lignocellulosic substrate to its constituent sugars. The Examples below include assays that may be used to identify optimum ratios and blends/compositions of enzymes with which to degrade lignocellulosic materials.

6.3.1 Background

[0122] The cell walls of higher plants are comprised of a variety of carbohydrate polymer (CP) components. These CP interact through covalent and non-covalent means, providing the structural integrity required to form rigid cell walls and resist turgor pressure in plants. The major CP found in plants is cellulose, which forms the structural backbone of the cell wall. During cellulose biosynthesis, chains of poly- β -1,4-D-glucose self associate through hydrogen bonding and hydrophobic interactions to form cellulose microfibrils, which further self-associate to form larger fibrils. Cellulose microfibrils are often irregular structurally and contain regions of varying crystallinity. The degree of crystallinity of cellulose fibrils depends on how tightly ordered the hydrogen bonding is between and among its component cellulose chains. Areas with less-ordered bonding, and therefore more accessible glucose chains, are referred to as amorphous regions.

[0123] The general model for cellulose depolymerization to glucose involves a minimum of three distinct enzymatic activities. Endoglucanases cleave cellulose chains internally to shorter chains in a process that increases the number of accessible ends, which are more susceptible to exoglucanase activity than the intact cellulose chains. These exoglucanases (e.g., cellobiohydrolases) are specific for either reducing ends or non-reducing ends, liberating, in most cases, cellobiose, the dimer of glucose.. The accumulating cellobiose is then subject to cleavage by cellobiases (e.g., β -1,4-glucosidases) to glucose.

[0124] Cellulose contains only anhydro-glucose. In contrast, hemicellulose contains a number of different sugar monomers. For instance, aside from glucose, sugar monomers in hemicellulose can also include xylose, mannose, galactose, rhamnose, and arabinose. Hemicelluloses mostly contain D-pentose sugars and occasionally small amounts of L-sugars. Xylose is typically present in the largest amount, but mannuronic acid and galacturonic acid also tend to be present. Hemicelluloses include xylan, glucuronoxylan, arabinoxylan, glucomannan, and xyloglucan.

[0125] The enzymes and multi-enzyme compositions of the disclosure are useful for saccharification of hemicellulose materials, including, 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.

[0126] Most biomass sources are rather complex, containing cellulose, hemicellulose, pectin, lignin, protein, and ash, among other components. Accordingly, in certain aspects, the present disclosure provides enzyme blends/compositions containing enzymes that impart a range or variety of substrate specificities when working together to degrade biomass into fermentable sugars in the most efficient manner. One example of a multi-enzyme blend/composition is a mixture of cellobiohydrolase(s), xylanase(s), endoglucanase(s), β -glucosidase(s), β -xylosidase(s), and, optionally, accessory proteins. The enzyme blend/composition is suitably a non-naturally occurring composition.

[0127] Accordingly, the disclosure provides enzyme blends/compositions (including products of manufacture) comprising a mixture of xylan-hydrolyzing, hemicellulose- and/or cellulose-hydrolyzing enzymes, which include at least one, several, or all of a cellulase, including a glucanase; a cellobiohydrolase; an L- α -arabinofuranosidase; a xylanase; a β -glucosidase; and a β -xylosidase. Preferably each of the enzyme blends/compositions of the disclosure comprises at least one enzyme of the disclosure. The present disclosure also provides enzyme blends/compositions that are non-naturally occurring compositions. As used herein, the term "enzyme blends/compositions" refers to:

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

[0128] The term "fermentation broth" as used herein refers to an enzyme preparation produced by fermentation that undergoes no or minimal recovery and/or purification subsequent to fermentation. For example, microbial cultures are grown to saturation, incubated under carbon-limiting conditions to allow protein synthesis (e.g., expression of enzymes). Then, once the enzyme(s) are secreted into the cell culture media, the fermentation broths can be used. The fermentation broths of the disclosure can contain unfractionated or fractionated contents of the fermentation materials derived at the end of the fermentation. For example, the fermentation broths are unfractionated and comprise the spent culture medium and cell debris present after the microbial cells (e.g., filamentous fungal cells) undergo a

fermentation process. The fermentation broth can suitably contain the spent cell culture media, extracellular enzymes, and live or killed microbial cells. Alternatively, the fermentation broths can be fractionated to remove the microbial cells. In those cases, the fermentation broths can, for example, comprise the spent cell culture media and the extracellular enzymes.

[0129] Any of the enzymes described specifically herein can be combined with any one or more of the enzymes described herein or with any other available and suitable enzymes, to produce a suitable multi-enzyme blend/composition. The disclosure is not restricted or limited to the specific exemplary combinations listed below.

6.3.2 Biomass

[0130] The disclosure provides methods and processes for biomass saccharification, using enzymes, enzyme blends/compositions of the disclosure. The term "biomass," as used herein, refers to any composition comprising cellulose and/or hemicellulose (optionally also lignin in lignocellulosic biomass materials). As used herein, biomass includes, without limitation, seeds, grains, tubers, plant waste or byproducts of food processing or industrial processing (e.g., stalks), corn (including, e.g., cobs, stover, and the like), grasses (including, e.g., Indian grass, such as *Sorghastrum nutans*; or, switchgrass, e.g., *Panicum* species, such as *Panicum virgatum*), wood (including, e.g., wood chips, processing waste), paper, pulp, and recycled paper (including, e.g., newspaper, printer paper, and the like). Other biomass materials include, without limitation, potatoes, soybean (e.g., rapeseed), barley, rye, oats, wheat, beets, and sugar cane bagasse.

[0131] The disclosure provides methods of saccharification comprising contacting a composition comprising a biomass material, for example, a material comprising xylan, hemicellulose, cellulose, and/or a fermentable sugar, with a polypeptide of the disclosure, or a polypeptide encoded by a nucleic acid of the disclosure, or any one of the enzyme blends/compositions, or products of manufacture of the disclosure.

[0132] The saccharified biomass (e.g., lignocellulosic material processed by enzymes of the disclosure) can be made into a number of bio-based products, *via* processes such as, e.g., microbial fermentation and/or chemical synthesis. As used herein, "microbial fermentation" refers to a process of growing and harvesting fermenting microorganisms under suitable conditions. The fermenting microorganism can be any microorganism suitable for use in a desired fermentation process for the production of bio-based products. Suitable fermenting microorganisms include, without limitation, filamentous fungi, yeast, and bacteria. The saccharified biomass can, for example, be made it into a fuel (e.g., a biofuel such as a bioethanol, biobutanol, biomethanol, a biopropanol, a biodiesel, a jet fuel, or the like) *via* fermentation and/or chemical synthesis. The saccharified biomass can, for example, also be made into a commodity chemical (e.g., ascorbic acid, isoprene, 1,3-propanediol), lipids, amino acids, proteins, and enzymes, *via* fermentation and/or chemical synthesis.

6.3.3. Pretreatment

[0133] Prior to saccharification, biomass (e.g., lignocellulosic material) is preferably subject to one or more pretreatment step(s) in order to render xylan, hemicellulose, cellulose and/or lignin material more accessible or susceptible to enzymes and thus more amenable to hydrolysis by the enzyme(s) and/or enzyme blends/compositions of the disclosure.

[0134] In an exemplary embodiment, the pretreatment entails subjecting biomass material to a catalyst comprising a dilute solution of a strong acid and a metal salt in a reactor. The biomass material can, e.g., be a raw material or a dried material. This pretreatment can lower the activation energy, or the temperature, of cellulose hydrolysis, ultimately allowing higher yields of fermentable sugars. See, e.g., U.S. Patent Nos. 6,660,506; 6,423,145.

[0135] Another exemplary pretreatment method entails hydrolyzing biomass by subjecting the biomass material to a first hydrolysis step in an aqueous medium at a temperature and a pressure chosen to effectuate primarily depolymerization of hemicellulose without achieving significant depolymerization of cellulose into glucose. This step yields a slurry in which the liquid aqueous phase contains dissolved monosaccharides resulting from depolymerization of hemicellulose, and a solid phase containing cellulose and lignin. The slurry is then subject to a second hydrolysis step under conditions that allow a major portion of the cellulose to be depolymerized, yielding a liquid aqueous phase containing dissolved/soluble depolymerization products of cellulose. See, e.g., U.S. Patent No. 5,536,325.

[0136] A further exemplary method involves processing a biomass material by one or more stages of dilute acid hydrolysis using about 0.4% to about 2% of a strong acid; followed by treating the unreacted solid lignocellulosic component of the acid hydrolyzed material with alkaline delignification. See, e.g., U.S. Patent No. 6,409,841.

[0137] Another exemplary pretreatment method comprises prehydrolyzing biomass (e.g., lignocellulosic materials) in a prehydrolysis reactor; adding an acidic liquid to the solid lignocellulosic material to make a mixture; heating the mixture to reaction temperature; maintaining reaction temperature for a period of time sufficient to fractionate the lignocellulosic material into a solubilized portion containing at least about 20% of the lignin from the lignocellulosic material, and a solid fraction containing cellulose; separating the solubilized portion from the solid fraction, and removing the solubilized portion while at or near reaction temperature; and recovering the solubilized portion. The cellulose in the solid fraction is rendered more amenable to enzymatic digestion. See, e.g., U.S. Patent No. 5,705,369.

[0138] Further pretreatment methods can involve the use of hydrogen peroxide H₂O₂. See Gould, 1984, Biotech, and Bioengr. 26:46-52.

[0139] Pretreatment can also comprise contacting a biomass material with stoichiometric

amounts of sodium hydroxide and ammonium hydroxide at a very low concentration. See Teixeira et al., 1999, Appl. Biochem. and Biotech. 77-79:19-34.

[0140] Pretreatment can also comprise contacting a lignocellulose with a chemical (e.g., a base, such as sodium carbonate or potassium hydroxide) at a pH of about 9 to about 14 at moderate temperature, pressure, and pH. See PCT Publication WO2004/081185.

[0141] Ammonia is used, for example, in a preferred pretreatment method. Such a pretreatment method comprises subjecting a biomass material to low ammonia concentration under conditions of high solids. See, e.g., U.S. Patent Publication No. 20070031918 and PCT publication WO 06110901.

6.3.4 Exemplary Enzyme Blends

[0142] The present disclosure provides enzyme blends/compositions comprising one or more enzymes of the disclosure. One or more enzymes of the enzyme blends/compositions can be produced by a recombinant host cell or a recombinant organism. The enzyme blends/compositions are suitably non-naturally occurring compositions.

[0143] An enzyme blend/composition of the disclosure can suitably comprise a first polypeptide having β -xylosidase activity, and further comprises 1, 2, 3, or 4 of a second polypeptide having β -xylosidase activity, one or more polypeptides having L- α -arabinofuranosidase activity, one or more polypeptides having xylanase activity, and one or more polypeptides having cellulase activity. The first polypeptide having β -xylosidase activity is, for example, an Fv3A, a Pf43A, an Fv43E, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, or an Fv39A polypeptide. The second polypeptide having β -xylosidase activity, if present, is, for example, different from the first polypeptide having β -xylosidase activity, and is suitably an Fv3A, a Pf43A, an Fv43E, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fv43D, an Fo43A, an Fv39A, or a *Trichoderma reesei* Bxl1 polypeptide. Each of the one or more polypeptides having L- α -arabinofuranosidase activity, if present, is, for example, an Af43A, a Pf51A, a Pa51A, an Fv43B, or an Fv51A polypeptide. Each of the one or more polypeptides having xylanase activity is, for example, a *Trichoderma reesei* Xyn, a *Trichoderma reesei* Xyn2, an AfuXyn2, or an AfuXyn5. Each of the one or more polypeptides having cellulase activity, if present, is, for example, an endoglucanase, for example, a *Trichoderma reesei* EG1 or EG2, a cellobiohydrolase, for example, a *Trichoderma reesei* CBH1 or CBH2, or a β -glucosidase, for example, a *Trichoderma reesei* Bgl1.

[0144] Another enzyme blend/composition of the disclosure can suitably comprise a first polypeptide having L- α -arabinofuranosidase activity, and further comprises 1, 2, 3, or 4 of a second polypeptide having L- α -arabinofuranosidase activity, one or more polypeptides having β -xylosidase activity, one or more polypeptides having xylanase activity, and/or one or more polypeptides having cellulase activity. The first L- α -arabinofuranosidase is an Af43A, a Pf51A, or an Fv51A polypeptide. The second L- α -arabinofuranosidase is different from the first L- α -

arabinofuranosidase, and is, for example, an Af43A, a Pf51A, a Pa51A, an Fv43B, or an Fv51A polypeptide. Each of the one or more β -xylosidases is, for example, an Fv3A, a Pf43A, an Fv43E, an Fv43A, an Fv43B, a Pa51A, an Fv43D, a Gz43A, an Fo43A, an Fv39A, or a *Trichoderma reesei* Bxl1 polypeptide. In certain embodiments, each of the one or more xylanases is a *Trichoderma reesei* Xyn3, a *Trichoderma reesei* Xyn2, an AfuXyn2, or an AfuXyn5 polypeptide. In certain embodiments, each of the one or more cellulases is independently an endoglucanase, for example, a *Trichoderma reesei* EG1 or EG2 polypeptide, a cellobiohydrolase, for example, a *Trichoderma reesei* CBH1 or CBH2 polypeptide, or a β -glucosidase, for example, a *Trichoderma reesei* Bgl1 polypeptide.

[0145] Xylanases: The xylanase(s) suitably constitutes about 0.05 wt.% to about 75 wt.% of the enzymes in an enzyme blend/ composition of the disclosure (*i.e.*, the percentage xylanase(s) is a weight percentage relative to the weight of all proteins in the composition) or a relative weight basis (*i.e.*, wherein the percentage xylanase(s) is a weight percentage relative to the combined weight of xylanases, β -xylosidases, cellulases, L- α -arabinofuranosidases, and accessory proteins). The ratio of any pair of proteins relative to each other can be readily calculated in the enzyme blends/compositions of the disclosure. Blends/compositions comprising enzymes in any weight ratio derivable from the weight percentages disclosed herein are contemplated. The xylanase content can be in a range wherein the lower limit is 0.05 wt.%, 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of enzymes in the enzyme blend/composition, and the upper limit is 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, or 75 wt.% of the total weight of enzymes in the enzyme blend/composition. The one or more xylanases in an enzyme blend or composition can represent, for example, 5 wt.% to 20 wt.%, 10 wt.% to 15 wt.%, or 15 wt.% to 25 wt.% of the total enzymes in the enzyme blend/composition. Exemplary suitable xylanases for inclusion in the enzyme blends/compositions of the disclosure are described in Section 6.3.6 below.

[0146] L- α -arabinofuranosidases: The L- α -arabinofuranosidase(s) suitably constitutes about 0.05 wt.% to about 75 wt.% of the total weight of all enzymes in a given enzyme blend/ composition (*i.e.*, wherein the percentage L- α -arabinofuranosidase(s) is a weight percentage relative to the weight of all proteins in the blend/composition) or a relative weight basis (*i.e.*, wherein the percentage L- α -arabinofuranosidase(s) is a weight percentage relative to the combined weight of xylanases, β -xylosidases, cellulases, L- α -arabinofuranosidases, and accessory proteins). The ratio of any pair of proteins relative to each other can be readily calculated based on the disclosure. Blends/compositions comprising enzymes in any weight ratio derivable from the weight percentages disclosed herein are contemplated. The L- α -arabinofuranosidase content can be in a range wherein the lower limit is 0.05 wt.%, 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of enzymes in the blend/composition, and the upper limit is 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.% or 75 wt.% of the total weight of enzymes in the blend/composition. For example, the one or more L- α -arabinofuranosidase(s)

can suitably represent 2 wt.% to 25 wt.%; 5 wt.% to 20 wt.%; or 5 wt.% to 10 wt.% of the total weight of enzymes in the blend/composition. Exemplary suitable L- α -arabinofuranosidase(s) for inclusion in the enzyme blends/compositions of the disclosure are described in Section 6.3.8 below.

[0147] β -Xylosidases: The β -xylosidase(s) suitably constitutes about 0.05 wt.% to about 75 wt.% of the total weight of enzymes in an enzyme blend/composition. The ratio of any pair of proteins relative to each other can be readily calculated based on the disclosure herein. Blends/compositions comprising enzymes in any weight ratio derivable from the weight percentages disclosed herein are contemplated. The β -xylosidase content can be in a range wherein the lower limit is about 0.05 wt.%, 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 40 wt.%, 45 wt.%, or 50 wt.% of the total weight of enzymes in the blend/composition, and the upper limit is about 10 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, 30 wt.%, 35 wt.%, 40 wt.%, 50 wt.%, 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, or 75 wt.% of the total weight of enzymes in the blend/composition. For example, the β -xylosidase(s) can represent about 0.05 wt.% to about 75 wt.% of the total weight of enzymes in the blend/composition. Also, the β -xylosidase(s) can represent 0.05 wt.% to about 70 wt.%, about 1 wt.% to about 65 wt.%, about 1 wt.% to about 60 wt.%, about 2 wt.% to about 55 wt.%, about 3 wt.% to about 50 wt.%, about 4 wt.% to about 45 wt.%, or about 5 wt.% to about 40 wt.% of the total weight of enzymes in the blend/composition. In yet a further example, the β -xylosidase(s) suitably represent 2 wt.% to 30 wt.%; 10 wt.% to 20 wt.%; or 5 wt.% to 10 wt.% of the total weight of enzymes in the blend/composition. Exemplary suitable β -xylosidase(s) are described in Section 6.3.7 below.

[0148] Cellulases: The cellulase(s) suitably constitutes about 0.05 wt.% to about 90 wt.% of the total weight of enzymes in an enzyme blend/composition. Ratio of any pair of proteins relative to each other can be readily calculated based on the disclosure herein. Blends/compositions comprising enzymes in any weight ratio derivable from the weight percentages disclosed herein are contemplated. The cellulase content can be in a range wherein the lower limit is about 0.05 wt.%, 5 wt.%, 10 wt.%, 20 wt.%, 30 wt.%, 40 wt.%, 50 wt.%, 60 wt.%, 70 wt.% of the total weight of enzymes in the blend/composition, and the upper limit is about 20 wt.%, 30 wt.%, 40 wt.%, 50 wt.%, 60 wt.%, 70 wt.%, 80 wt.%, or 90 wt.% of the total weight of enzymes in the blend/composition. For example, the cellulase(s) suitably represents 30 wt.% to 80 wt.%, 50 wt.% to 70 wt.%, or 40 wt.% to 60 wt.% of the total weight of enzymes in the blend/composition. Exemplary suitable cellulases are described in Section 6.3.5 below. The cellulase components in an enzyme blend/composition of the disclosure are suitably capable of achieving at least about 0.005 fraction product per mg protein per gram of phosphoric acid swollen cellulose (PASC) as determined by a calcofluor assay. For example, the cellulase components in a blend/composition of the disclosure are capable of achieving a range of fraction product per mg protein per gram of PASC, wherein the lower limit of the range is about 0.005, 0.01, 0.015, 0.02, 0.03, 0.04, 0.05, 0.06, 0.075, or 0.1, and wherein the upper limit of the range is, 0.03, 0.04, 0.05, 0.06, 0.075, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, or 0.7. The cellulase components in a blend/composition of the disclosure can, for example, achieve 0.00005 - 0.0001, 0.0005 - 0.001, 0.001 - 0.005, 0.005 - 0.03, 0.01 - 0.06, 0.02 - 0.04, 0.01 - 0.03,

0.02-0.05, 0.02-0.04, 0.01-0.05, 0.015-0.035, or 0.015-0.075 product fraction product per mg protein per gram PASC as determined by a calcofluor assay. The cellulase can be, for example, a whole cellulase. The cellulase can also, for example, suitably be enriched with a β -glucosidase.

[0149] Accessory Proteins: The enzyme blend/composition may suitably further comprise one or more accessory proteins. The accessory protein content of an enzyme blend/composition can range from about 0 wt.% to about 60 wt.% of the total weight of proteins in an enzyme blend/composition. Ratio of any pair of proteins relative to each other can be readily determined based on the disclosure herein. Blends/compositions comprising enzymes in any weight ratio derivable from the weight percentages disclosed herein are contemplated. The accessory protein content can be in a range wherein the lower limit is 0 wt.%, 1 wt.%, 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 10 wt.%, 12 wt.%, 15 wt.%, 20 wt.%, 25 wt.%, or 35 wt.% of the total weight of proteins in the enzyme blend/composition, and the upper limit is 2 wt.%, 3 wt.%, 4 wt.%, 5 wt.%, 6 wt.%, 7 wt.%, 8 wt.%, 9 wt.%, 10 wt.%, 11 wt.%, 12 wt.%, 13 wt.%, 14 wt.%, 15 wt.%, 20 wt.%, 30 wt.%, 40 wt.%, 50 wt.%, or 60 wt.% of the total weight of proteins in the enzyme blend/composition. For example, the accessory protein(s) can suitably represent 0 wt.% to 2 wt.%, 5 wt.% to 10 wt.%, 20 wt.% to 50 wt.%, or 2 wt.% to 5 wt.% in the enzyme blend/composition. Exemplary suitable accessory proteins for inclusion in the enzyme blends/compositions of the disclosure are described in Section 6.3.9 below.

[0150] The present disclosure provides a first enzyme blend/composition for lignocellulose saccharification comprising:

1. (1) about 30 wt.% to about 80 wt.% (e.g., 30 wt.% to 80 wt.%, 35 wt.% to 75 wt.%, 40 wt.% to 70 wt.%, 40 wt.% to 60 wt.%, 50 wt.% to 70 wt.%, etc.) cellulase(s), e.g., whole cellulase or β -glucosidase enriched whole cellulase;
2. (2) about 3 wt.% to about 50 wt.% (e.g., 5 wt.% to 40 wt.%, 10 wt.% to 30 wt.%, 5 wt.% to 20 wt.%, 10 wt.% to 15 wt.%, 15 wt.% to 25 wt.%, etc.) xylanase(s), e.g., a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, an AfuXyn5, or a mixture of two or more of the foregoing enzymes;
3. (3) about 2 wt.% to about 40 wt.% (e.g., 2 wt.% to 35 wt.%, 5 wt.% to 30 wt.%, 10 wt.% to 25 wt.%, 2 wt.% to 30 wt.%, 10 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc.) β -xylosidase(s), e.g., an Fv3A, a Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fv43A, an Fv43B, a Pa51A, an Fo43A, a Gz43A, a *Trichoderma reesei* Bxl1, or a mixture of two or more of the foregoing enzymes;
4. (4) about 2 wt.% to about 40 wt.% (e.g., 2 wt.% to 35 wt.%, 5 wt.% to 30 wt.%, 10 wt.% to 25 wt.%, 2 wt.% to 25 wt.%, 5 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc.) L- α -arabinofuranosidase(s), e.g., an Af43A, an Fv43B, a Pa51A, a Pf51A, an Fv51A, or a mixture of two or more of the foregoing enzymes; and
5. (5) about 0 wt.% to about 50 wt.% (2 wt.% to 40 wt.%, 5 wt.% to 30 wt.%, 10 wt.% to 25 wt.%, 0 wt.% to 2 wt.%, 5 wt.% to 10 wt.%, 20 wt.% to 50 wt.%, 2 wt.% to 5 wt.%, etc.) accessory protein(s).

[0151] The present disclosure provides a second enzyme blend/composition for lignocellulose saccharification comprising:

1. (1) about 30 wt.% to about 80 wt.% (e.g., 30 wt.% to 80 wt.%, 35 wt.% to 75 wt.%, 40 wt.% to 70 wt.%, 40 wt.% to 60 wt.%, 50 wt.% to 70 wt.%, etc.) cellulase(s), e.g., whole cellulase or β -glucosidase enriched whole cellulase, or about 2 wt.% to about 10 wt.% (e.g., 2 wt.% to 8 wt.%, 4 wt.% to 6 wt.%, 2 wt.% to 4 wt.%, 6 wt.% to 8 wt.%, 8 wt.% to 10 wt.%, 2 wt.% to 10 wt.%, etc) β -glucosidase(s), e.g., a *Trichoderma reesei* Bgl1;
2. (2) about 3 wt.% to about 50 wt.% (e.g., 5 wt.% to 40 wt.%, 10 wt.% to 30 wt.%, 5 wt.% to 20 wt.%, 10 wt.% to 15 wt.%, 15 wt.% to 25 wt.%, etc.) xylanase(s), e.g., a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, an AfuXyn5, or a mixture of two or more of the foregoing enzymes;
3. (3) about 2 wt.% to about 40 wt.% (e.g., 2 wt.% to 35 wt.%, 5 wt.% to 30 wt.%, 10 wt.% to 25 wt.%, 2 wt.% to 30 wt.%, 10 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc.) of at least two β -xylosidase(s), wherein at least one β -xylosidase is selected from Group 1 and at least one β -xylosidase is selected from Group 2;

wherein:

Group 1: an Fv3A, an Fv43A, or a mixture thereof;

Group 2: an Fv43D, a Pa51A, a Gz43A, a *Trichoderma reesei* Bxl1, a Pf43A, an Fv43E, an Fv39A, an Fo43A, an Fv43B, or a mixture of two or more of the foregoing enzymes;

4. (4) about 2 wt.% to about 25 wt.% (e.g., 2 wt.% to 25 wt.%, 5 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc) L- α -arabinofuranosidase(s), e.g., an Af43A, an Fv43B, a Pa51A, a Pf51A, Fv51A, or a mixture of two or more of the foregoing enzymes; and
5. (5) about 0 wt.% to about 50 wt.% (2 wt.% to 40 wt.%, 5 wt.% to 30 wt.%, 10 wt.% to 25 wt.%, 0 wt.% to 2 wt.%, 5 wt.% to 10 wt.%, 20 wt.% to 50 wt.%, 2 wt.% to 5 wt.%, etc) accessory protein(s).

[0152] The present disclosure provides a third enzyme blend/composition for lignocellulose saccharification comprising:

1. (1) about 3 wt.% to about 50 wt.% (e.g., 5 wt.% to 40 wt.%, 10 wt.% to 30 wt.%, 5 wt.% to 20 wt.%, 10 wt.% to 15 wt.%, 15 wt.% to 25 wt.%, etc.) xylanase(s), e.g., a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, an AfuXyn5, or a mixture of two or more of the foregoing enzymes; and
 2. (2) about 2 wt.% to 40 wt.% (e.g., 2 wt.% to 35 wt.%, 5 wt.% to 30 wt.%, 10 wt.% to 25 wt.%, 2 wt.% to 30 wt.%, 10 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc.) of at least two β -xylosidase(s), wherein at least one β -xylosidase is selected from Group 1 and at least one β -xylosidase is selected from Group 2;
- wherein:

Group 1: an Fv3A, an Fv43A, or a mixture thereof;

Group 2: an Fv43D, a Pa51A, a Gz43A, a *Trichoderma reesei* Bxl1, a Pf43A, an Fv43E, an Fv39A, an Fo43A, an Fv43B, or a mixture of two or more of the foregoing enzymes.

[0153] The present disclosure provides a fourth enzyme blend/composition for lignocellulose saccharification comprising:

1. (1) about 5 wt.% to about 25 wt.% (e.g., 2 wt.% to 25 wt.%, 5 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc) xylanase(s), e.g., a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, an AfuXyn5, or a mixture of two or more of the foregoing enzymes;
2. (2) about 2 wt.% to about 30 wt.% (e.g., 2 wt.% to 30 wt.%, 10 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc.) β -xylosidase(s), e.g., an Fv3A, a Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, a *Trichoderma reesei* Bxl1, or a mixture of two or more of the foregoing enzymes; and
3. (3) about 2 wt.% to about 50 wt.% (e.g., 2 wt.% to 5 wt.%, 5 wt.% to 45 wt.%, 10 wt.% to 40 wt.%, 15 wt.% to 30 wt.%, 10 wt.% to 25 wt.%, 5 wt.% to 20 wt.%, 15 wt.% to 40 wt.%, etc) β -glucosidase(s), e.g., a *Trichoderma reesei* Bgl1.

[0154] The present disclosure provides a fifth enzyme blend/composition for lignocellulose saccharification comprising:

1. (1) about 5 wt.% to about 25 wt.% (e.g., 2 wt.% to 25 wt.%, 5 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc) xylanase(s), e.g., a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, an AfuXyn5, or a mixture of two or more of the foregoing enzymes; and
2. (2) about 2 wt.% to about 30 wt.% (e.g., 2 wt.% to 30 wt.%, 10 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc.) β -xylosidase(s), e.g., an Fv3A, a Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, a *Trichoderma reesei* Bxl1, or a mixture of two or more of the foregoing enzymes.

[0155] The present disclosure provides a sixth enzyme blend/composition for lignocellulose saccharification comprising:

1. (1) about 2 wt.% to about 50 wt.% (e.g., 2 wt.% to 5 wt.%, 5 wt.% to 45 wt.%, 10 wt.% to 40 wt.%, 15 wt.% to 30 wt.%, 10 wt.% to 25 wt.%, 5 wt.% to 20 wt.%, 15 wt.% to 40 wt.%, etc) β -glucosidase(s), e.g., a Bgl1;
2. (2) about 3 wt.% to about 50 wt.% (e.g., 5 wt.% to 40 wt.%, 10 wt.% to 30 wt.%, 5 wt.% to 20 wt.%, 10 wt.% to 15 wt.%, 15 wt.% to 25 wt.%, etc.) xylanase(s), e.g., a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, an AfuXyn5, or a mixture of two or more of the foregoing enzymes;

3. (3) about 2 wt.% to 40 wt.% (e.g., 2 wt.% to 35 wt.%, 5 wt.% to 30 wt.%, 10 wt.% to 25 wt.%, 2 wt.% to 30 wt.%, 10 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc.) β -xylosidase(s), e.g., an Fv3A, a Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fv43A, an Fv43B, a Pa51A, a Gz43A, an Fo43A, a *Trichoderma reesei* Bxl1, or a mixture of two or more of the foregoing enzymes; and
4. (4) about 2 wt.% to about 25 wt.% (e.g., 2 wt.% to 25 wt.%, 5 wt.% to 20 wt.%, 5 wt.% to 10 wt.%, etc) L- α -arabinofuranosidase(s), e.g., an Af43A, an Fv43B, a Pa51A, a Pf51A, Fv51A, or a mixture of two or more of the foregoing enzymes.

[0156] The sixth enzyme blend/composition for lignocellulose saccharification above can, for example, comprise about 2 wt.% to about 40 wt.% of at least two β -xylosidase, wherein at least one β -xylosidase is selected from Group 1 and at least one β -xylosidase is selected from Group 2; wherein:

Group 1: an Fv3A, an Fv43A, or a mixture thereof;

Group 2: an Fv43D, a Pa51A, a Gz43A, a *Trichoderma reesei* Bxl1, a Pf43A, an Fv43E, an Fv39A, an Fo43A, an Fv43B, or a mixture of two or more of the foregoing enzymes.

[0157] Where an enzyme blend/composition of the disclosure contains both Group 1 and a Group 2 β -xylosidases, the ratio of Group 1 to Group 2 β -xylosidases is preferably 1:10 to 10:1. For example, the ratio is suitably 1:2 to 2:1, 2:5 to 5:2, 3:8 to 8:3, 1:4 to 4:1, 1:5 to 5:1, 1:7 to 7:1, or any range between any pair the foregoing endpoints (e.g., 1:10 to 2:1, 4:1 to 2:5, 3:8 to 5:1, etc.). A particular example of a suitable ratio is approximately 1:1.

[0158] Where an enzyme blend/composition of the disclosure contains an Fv43A as a β -xylosidase, the blend/composition can further contain Fv43B as an L- α -arabinofuranosidase.

[0159] An enzyme blend/composition of the disclosure is, for example, suitably part of a saccharification reaction mixture containing biomass in addition to the components of the enzyme blend/composition. For example, the saccharification reaction mixture can be characterized by 1, 2, 3 or all 4 of the following features:

1. (1) the total weight of xylanase(s) per kg of hemicellulase in said saccharification reaction mixture is in a range in which the lower limit is about 0.5 g, 1 g, 2 g, 3 g, 4 g, 5 g, 7 g, or 10 g, and the upper limit is independently about 5 g, 7 g, 10 g, 15 g, 20 g, 30 g, or 40 g; for example, the total weight of xylanase(s) per kg of hemicellulase in the reaction mixture can be 0.5 g to 40 g, 0.5g to 30 g, 0.5 g to 20 g, 0.5 to 10 g, 0.5 to 5 g, 1 g to 40 g, 2 g to 40 g, 3 g to 40 g, 5 g to 40 g, 7 g to 30 g, 10 g to 30 g, 5 g to 20 g, or 5 g to 30 g.
- (ii) the total weight of β -xylosidase(s) per kg of hemicellulase in said saccharification

reaction mixture is in a range in which the lower limit is about 0.5 g, 1 g, 2 g, 3 g, 4 g, 5 g, 7 g, or 10 g and the upper limit is independently about 5 g, 7 g, 10 g, 15 g, 20 g, 30 g, 40 g, or 50 g; for example, the total weight of β -xylosidase(s) per kg of hemicellulase in the reaction mixture can be 0.5 g to 40 g, 0.5 to 50 g, 0.5 g to 30 g, 0.5 g to 20 g, 0.5 g to 10 g, 0.5 g to 5g, 1 g to 40 g, 2 g to 40 g, 3 g to 40 g, 5 g to 40 g, 7 g to 30 g, 10 g, to 30 g, 5 g to 30 g, 5 g to 20 g.

(iii) the total weight of L- α -arabinofuranosidase(s) per kg of hemicellulase in said saccharification reaction mixture is in a range in which the lower limit is about 0.2 g, 0.5 g, 1 g, 1.5 g, 2 g, 2.5 g, 3 g, 4 g, or 5 g and the upper limit is independently about 2 g, 3 g, 4 g, 5 g, 7 g, 10 g, 15 g, or 20 g; for example, the total weight of L- α -arabinofuranosidase(s) per kg of hemicellulase in the reaction mixture can be 0.2 g to 20 g, 0.5 g to 20 g, 1 g to 20 g, 2 g to 20 g, 2.5 g to 20 g, 3 g to 15 g, 4 g to 20 g, 5 g to 15 g, 5 g to 10 g, 5 g to 20 g, or 2.5 g to 15 g.

(iv) the total weight of cellulase(s) per kg of cellulase in said saccharification reaction mixture is in a range in which the lower limit is about 1 g, 3 g, 5 g, 7 g, 10 g, 12 g, 15 g, 18 g, or 20g, and the upper limit is independently about 10 g, 15 g, 18 g, 20 g, 25 g, 30g, 50 g, 75 g, or 100g; for example, the total weight of cellulase(s) per kg of cellulase in said reaction mixture can be 1 g to 100 g, 3 g to 100 g, 5 g to 100 g, 7 g to 100 g, 12 g to 100 g, 15 g to 100 g, 18 g to 100 g, 3 g to 75 g, 5 g to 50 g, 7 g to 75 g, 10 g to 75 g, 10 g to 50 g, 12 g to 75 g, 12 g to 50 g, 15 g to 75 g, 15 g to 50 g, 18 g to 30 g, 18 g to 75 g.

6.3.5. Cellulases

[0160] The enzyme blends/compositions of the disclosure can comprise one or more cellulases. Cellulases are enzymes that hydrolyze cellulose (β -1,4-glucan or β D-glucosidic linkages) resulting in the formation of glucose, cellobiose, cellooligosaccharides, and the like. Cellulases have been traditionally divided into three major classes: endoglucanases (EC 3.2.1.4) ("EG"), exoglucanases or cellobiohydrolases (EC 3.2.1.91) ("CBH") and β -glucosidases (β -D-glucoside glucohydrolase; EC 3.2.1.21) ("BG") (Knowles et al., 1987, Trends in Biotechnology 5(9):255-261; Shulein, 1988, Methods in Enzymology, 160:234-242). Endoglucanases act mainly on the amorphous parts of the cellulose fiber, whereas cellobiohydrolases are also able to degrade crystalline cellulose.

[0161] Cellulases for use in accordance with the methods and compositions of the disclosure can be obtained from, or produced recombinantly from, *inter alia*, one or more of the following organisms: *Crinipellis scapella*, *Macrophomina phaseolina*, *Myceliophthora thermophila*, *Sordaria fimicola*, *Volutella colletotrichoides*, *Thielavia terrestris*, *Acremonium sp.*, *Exidia glandulosa*, *Fomes fomentarius*, *Spongipellis sp.*, *Rhizophlyctis rosea*, *Rhizomucor pusillus*, *Phycomyces niteus*, *Chaetostylum fresenii*, *Diplodia gossypina*, *Ulospora bilgramii*, *Saccobolus*

dilutellus, *Penicillium verruculosum*, *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.

[0162] For example, a cellulase for use in the method and/or composition of the disclosure is a whole cellulase and/or is capable of achieving at least 0.1 (e.g. 0.1 to 0.4) fraction product as determined by the calcofluor assay described in Section 7.1.10 below.

6.3.5.1 β -Glucosidase

[0163] The enzyme blends/compositions of the disclosure can optionally comprise one or more β -glucosidases. The term " β -glucosidase" as used herein refers to a β -D-glucoside glucohydrolase classified as EC 3.2.1.21, and/or members of certain GH families, including, without limitation, members of GH families 1, 3, 9 or 48, which catalyze the hydrolysis of cellobiose to release β -D-glucose.

[0164] Suitable β -glucosidase can be obtained from a number of microorganisms, by recombinant means, or be purchased from commercial sources. Examples of β -glucosidases from microorganisms include, without limitation, ones from bacteria and fungi. For example, a β -glucosidase of the present disclosure is suitably obtained from a filamentous fungus.

[0165] The β -glucosidases can be obtained, or produced recombinantly, from, *inter alia*, *Aspergillus aculeatus* (Kawaguchi et al. Gene 1996, 173: 287-288), *Aspergillus kawachi* (Iwashita et al. Appl. Environ. Microbiol. 1999, 65: 5546-5553), *Aspergillus oryzae* (WO 2002/095014), *Cellulomonas biazotea* (Wong et al. Gene, 1998, 207:79-86), *Penicillium funiculosum* (WO 2004/078919), *Saccharomycopsis fibuligera* (Machida et al. Appl. Environ. Microbiol. 1988, 54: 3147-3155), *Schizosaccharomyces pombe* (Wood et al. Nature 2002, 415: 871-880), or *Trichoderma reesei* (e.g., β -glucosidase 1 (U.S. Patent No. 6,022,725), β -glucosidase 3 (U.S. Patent No. 6,982,159), β -glucosidase 4 (U.S. Patent No. 7,045,332), β -glucosidase 5 (US Patent No. 7,005,289), β -glucosidase 6 (U.S. Publication No. 20060258554), β -glucosidase 7 (U.S. Publication No. 20060258554)).

[0166] The β -glucosidase can be produced by expressing an endogenous or exogenous gene encoding a β -glucosidase. For example, β -glucosidase can be secreted into the extracellular space e.g., by Gram-positive organisms (e.g., *Bacillus* or *Actinomycetes*), or eukaryotic hosts (e.g., *Trichoderma*, *Aspergillus*, *Saccharomyces*, or *Pichia*). The β -glucosidase can be, in some

circumstances, overexpressed or underexpressed.

[0167] The β -glucosidase can also be obtained from commercial sources. Examples of commercial β -glucosidase preparation suitable for use in the present disclosure include, for example, *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 from Megazyme (Megazyme International Ireland Ltd., Ireland.).

[0168] Moreover, the β -glucosidase can be a component of a whole cellulase, as described in Section 6.3.5.4 below.

[0169] β -glucosidase activity can be determined by a number of suitable means known in the art, such as the assay described by Chen et al., in Biochimica et Biophysica Acta 1992, 121:54-60, wherein 1 pNPG denotes 1 μ mol of Nitrophenol liberated from 4-nitrophenyl-p-D-glucopyranoside in 10 min at 50°C (122°F) and pH 4.8.

6.3.5.2 Endoglucanases

[0170] The enzyme blends/compositions of the disclosure optionally comprise one or more endoglucanase. Any endoglucanase (EC 3.2.1.4) can be used in the methods and compositions of the present disclosure. An endoglucanase can be produced by expressing an endogenous or exogenous endoglucanase gene. The endoglucanase can be, in some circumstances, overexpressed or underexpressed.

[0171] For example, *Trichoderma reesei* EG1 (Penttila et al., Gene 1986, 63:103-112) and/or EG2 (Saloheimo et al., Gene 1988, 63:11-21) are suitably used in the methods and compositions of the present disclosure.

[0172] A thermostable *Thielavia terrestris* endoglucanase (Kvesitadaze et al., Applied Biochem. Biotech. 1995, 50:137-143) is, in another example, used in the methods and compositions of the present disclosure. Moreover, a *Trichoderma reesei* EG3 (Okada et al. Appl. Environ. Microbiol. 1988, 64:555-563), EG4 (Saloheimo et al. Eur. J. Biochem. 1997, 249:584-591), EG5 (Saloheimo et al. Molecular Microbiology 1994, 13:219-228), EG6 (U.S. Patent Publication No. 20070213249), or EG7 (U.S. Patent Publication No. 20090170181), an *Acidothrmus cellulolyticus* EI endoglucanase (U.S. Pat. No. 5,536,655), a *Humicola insolens* endoglucanase V (EGV) (Protein Data Bank entry 4ENG), a *Staphylotrichum coccosporum* endoglucanase (U.S. Patent Publication No. 20070111278), an *Aspergillus aculeatus* endoglucanase F1-CMC (Ooi et al. Nucleic Acid Res. 1990, 18:5884), an *Aspergillus kawachii* IFO 4308 endoglucanase CMCase-1 (Sakamoto et al. Curr. Genet. 1995, 27:435-439), an *Erwinia carotovora* (Saarilahti et al. Gene 1990, 90:9-14); or an *Acremonium thermophilum* ALKO4245 endoglucanase (U.S. Patent Publication No. 20070148732) can also be used. Additional suitable endoglucanases are described in, e.g., WO 91/17243, WO 91/17244, WO

91/10732, U.S. Patent No. 6,001,639.

6.3.5.3 Cellobiohydrolases

[0173] Any cellobiohydrolase (EC 3.2.1.91) ("CBH") can be optionally used in the methods and blends/compositions of the present disclosure. The cellobiohydrolase can be produced by expressing an endogeneous or exogeneous cellobiohydrolase gene. The cellobiohydrolase can be, in some circumstances, overexpressed or under expressed.

[0174] For example, *Trichoderma reesei* CBHI (Shoemaker et al. Bio/Technology 1983, 1:691-696) and/or CBHII (Teeri et al. Bio/Technology 1983, 1:696-699) can be suitably used in the methods and blends/compositions of the present disclosure.

[0175] Suitable CBHs can be selected from an *Agaricus bisporus* CBH1 (Swiss Prot Accession No. Q92400), an *Aspergillus aculeatus* CBH1 (Swiss Prot Accession No. O59843), an *Aspergillus nidulans* CBHA (GenBank Accession No. AF420019) or CBHB (GenBank Accession No. AF420020), an *Aspergillus niger* CBHA (GenBank Accession No. AF156268) or CBHB (GenBank Accession No. AF156269), a *Claviceps purpurea* CBH1 (Swiss Prot Accession No. O00082), a *Cochliobolus carbonarum* CBH1 (Swiss Prot Accession No. Q00328), a *Cryphonectria parasitica* CBH1 (Swiss Prot Accession No. Q00548), a *Fusarium oxysporum* CBH1 (Cel7A) (Swiss Prot Accession No. P46238), a *Humicola grisea* CBH1.2 (GenBank Accession No. U50594), a *Humicola grisea* var. *thermoidea* CBH1 (GenBank Accession No. D63515) a CBH1.2 (GenBank Accession No. AF123441), or an exo1 (GenBank Accession No. AB003105), a *Melanocarpus albomyces* Cel7B (GenBank Accession No. AJ515705), a *Neurospora crassa* CBHI (GenBank Accession No. X77778), a *Penicillium funiculosum* CBHI (Cel7A) (U.S. Patent Publication No. 20070148730), a *Penicillium janthinellum* CBHI (GenBank Accession No. S56178), a *Phanerochaete chrysosporium* CBH (GenBank Accession No. M22220), or a CBHI-2 (Cel7D) (GenBank Accession No. L22656), a *Talaromyces emersonii* CBH1A (GenBank Accession No. AF439935), a *Trichoderma viride* CBH1 (GenBank Accession No. X53931), or a *Volvariella volvacea* V14 CBH1 (GenBank Accession No. AF156693).

6.3.5.4 Whole Cellulases

[0176] An enzyme blend/composition of the disclosure can further comprise a whole cellulase. As used herein, a "whole cellulase" refers to either a naturally occurring or a non-naturally occurring cellulase-containing composition comprising at least 3 different enzyme types: (1) an endoglucanase, (2) a cellobiohydrolase, and (3) a β -glucosidase, or comprising at least 3 different enzymatic activities: (1) an endoglucanase activity, which catalyzes the cleavage of internal β -1,4 linkages, resulting in shorter glucooligosaccharides, (2) a cellobiohydrolase activity, which catalyzes an "exo"-type release of cellobiose units (β -1,4 glucose-glucose disaccharide), and (3) a β -glucosidase activity, which catalyzes the release of glucose

monomer from short celooligosaccharides (e.g., cellobiose).

[0177] A "naturally occurring cellulase-containing" composition is one produced by a naturally occurring source, which comprises one or more cellobiohydrolase-type, one or more endoglucanase-type, and one or more β -glucosidase-type components or activities, wherein each of these components or activities is found at the ratio and level produced in nature, untouched by the human hand. Accordingly, a naturally occurring cellulase-containing composition is, for example, one that is produced by an organism unmodified with respect to the cellulolytic enzymes such that the ratio or levels of the component enzymes are unaltered from that produced by the native organism in nature. A "non-naturally occurring cellulase-containing composition" refers to a composition produced by: (1) combining component cellulolytic enzymes either in a naturally occurring ratio or a non-naturally occurring, *i.e.*, altered, ratio; or (2) modifying an organism to overexpress or underexpress one or more cellulolytic enzymes; or (3) modifying an organism such that at least one cellulolytic enzyme is deleted. A "non-naturally occurring cellulase containing" composition can also refer to a composition resulting from adjusting the culture conditions for a naturally-occurring organism, such that the naturally-occurring organism grows under a non-native condition, and produces an altered level or ratio of enzymes. Accordingly, in some embodiments, the whole cellulase preparation of the present disclosure can have one or more EGs and/or CBHs and/or β -glucosidases deleted and/or overexpressed.

[0178] 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 *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. Biotechnology, 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.

[0179] 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.

[0180] 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).

[0181] 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.

[0182] 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

[0183] 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 kills the cells after fermentation. The cells can, for example, be lysed or permeabilized using methods known in the art.

[0184] 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).

[0185] 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.

6.3.6 Xylanases

[0186] The enzyme blends/compositions of the disclosure, for example, can, comprise one or more Group A xylanases, which may be a *Trichoderma reesei* Xyn2, a *Trichoderma reesei* Xyn3, an AfuXyn2, or an AfuXyn5. Suitable *Trichoderma reesei* Xyn2, *Trichoderma reesei* Xyn3, AfuXyn2, or AfuXyn5 polypeptides are described in Section 6.1 above.

[0187] The enzyme blends/compositions of the disclosure optionally comprise one or more xylanases in addition to or in place of the one or more Group A xylanases. Any xylanase (EC 3.2.1.8) can be used as the additional one or more xylanases. Suitable xylanases include, e.g., 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, Journal of Fermentation and Bioengineering 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), a *Bacillus subtilis* xylanase (Bernier et al. 1983, Gene 26(1):59-65), a *Cellulomonas fimi* xylanase (Clarke et al., 1996, FEMS Microbiology Letters 139:27-35), a *Pseudomonas fluorescens* xylanase (Gilbert et al. 1988, Journal of General Microbiology 134:3239-3247), a *Clostridium thermocellum* xylanase (Dominguez et al., 1995, Nature Structural Biology 2:569-576), a *Bacillus pumilus* xylanase (Nuyens et al. Applied Microbiology and Biotechnology 2001, 56:431-434; Yang et al. 1998, 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).

[0188] The xylanase can be produced by expressing an endogenous or exogenous gene encoding a xylanase. The xylanase can be, in some circumstances, overexpressed or underexpressed.

6.3.7 β -Xylosidases

[0189] The enzyme blends/compositions of the disclosure, for example, can suitably comprise one or more β -xylosidases. For example, the β -xylosidase is a Group 1 β -xylosidase enzyme (e.g., an Fv3A or an Fv43A) or a Group 2 β -xylosidase enzyme (e.g., a Pf43A, an Fv43D, an Fv39A, an Fv43E, an Fo43A, an Fv43B, a Pa51A, a Gz43A, or a *Trichoderma reesei* Bxl1). These polypeptides are described in Section 0 above. For example, an enzyme blend/composition of the disclosure can suitably comprise one or more Group 1 β -xylosidases and one or more Group 2 β -xylosidases.

[0190] The enzyme blends/compositions of the disclosure can optionally comprise one or more β -xylosidases, in addition to or in place of the Group 1 and/or Group 2 β -xylosidases above. Any β -xylosidase (EC 3.2.1.37) can be used as the additional β -xylosidases. Suitable β -xylosidases include, for example, a *Talaromyces emersonii* Bxl1 (Reen et al. 2003, Biochem Biophys Res Commun. 305(3):579-85), a *Geobacillus stearothermophilus* β -xylosidases (Shallom et al. 2005, Biochemistry 44:387-397), a *Scytalidium thermophilum* β -xylosidases (Zanoelo et al. 2004, J. Ind. Microbiol. Biotechnol. 31:170-176), a *Trichoderma lignorum* β -xylosidases (Schmidt, 1998, Methods Enzymol. 160:662-671), an *Aspergillus awamori* β -xylosidases (Kurakake et al. 2005, Biochim. Biophys. Acta 1726:272-279), an *Aspergillus versicolor* β -xylosidases (Andrade et al. 2004, Process Biochem. 39:1931-1938), a *Streptomyces sp.* β -xylosidases (Pinphanichakarn et al. 2004, World J. Microbiol. Biotechnol. 20:727-733), a *Thermotoga maritima* β -xylosidases (Xue and Shao, 2004, Biotechnol. Lett. 26:1511-1515), a *Trichoderma sp.* SY β -xylosidases (Kim et al. 2004, J. Microbiol. Biotechnol. 14:643-645), an *Aspergillus niger* β -xylosidases (Oguntimein and Reilly, 1980, Biotechnol. Bioeng. 22:1143-1154), or a *Penicillium wortmanni* β -xylosidases (Matsuo et al. 1987, Agric. Biol. Chem. 51:2367-2379).

[0191] The β -xylosidase can be produced by expressing an endogenous or exogenous gene encoding a β -xylosidase. The β -xylosidase can be, in some circumstances, overexpressed or underexpressed.

6.3.8 L- α -Arabinofuranosidases

[0192] The enzyme blends/compositions of the disclosure can, for example, suitably comprise one or more L- α -arabinofuranosidases. The L- α -arabinofuranosidase is, for example, an Af43A, an Fv43B, a Pf51A, a Pa51A, or an Fv51A. Af43A, Fv43B, Pf51A, Pa51A, and Fv51A polypeptides are described in Section 6.1 above.

[0193] The enzyme blends/compositions of the disclosure optionally comprise one or more L- α -arabinofuranosidases in addition to or in place of the foregoing L- α -arabinofuranosidases. L- α -arabinofuranosidases (EC 3.2.1.55) from any suitable organism can be used as the additional L- α -arabinofuranosidases. Suitable L- α -arabinofuranosidases include, e.g., an L- α -arabinofuranosidases of *Aspergillus oryzae* (Numan & Bhosle, J. Ind. Microbiol. Biotechnol.

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

[0194] The L- α -arabinofuranosidase can be produced by expressing an endogenous or exogenous gene encoding an L- α -arabinofuranosidase. The L- α -arabinofuranosidase can be, in some circumstances, overexpressed or underexpressed.

6.3.9 Accessory proteins

[0195] 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 Further Applications

[0196] In addition to saccharification of biomass, the enzymes and/or enzyme blends/compositions of the disclosure can be used in industrial, agricultural, food and feed, as well as food and feed supplement processing processes. Exemplary applications are described below.

6.4.1 Wood, Paper and Pulp Treatments

[0197] The enzymes, enzyme blends/compositions, and methods of the disclosure can be used in wood, wood product, wood waste or by-product, paper, paper product, paper or wood pulp, Kraft pulp, or wood or paper recycling treatment or industrial process. These processes include, e.g., treatments of wood, wood pulp, paper waste, paper, or pulp, or deinking of wood or paper. The enzymes, enzyme blends/compositions of the disclosure can be, for example, used to treat/pretreat paper pulp, or recycled paper or paper pulp, and the like. The enzymes, enzyme blends/compositions of the disclosure can be used to increase the "brightness" of the paper when they are included in the paper, pulp, recycled paper or paper pulp treatment/pretreatment. It can be appreciated that the higher the grade of paper, the greater the brightness; the brightness can impact the scan capability of optical scanning equipment. As such, the enzymes, enzyme blends/compositions, and methods/processes can be used to make high grade, "bright" papers, including inkjet, laser and photo printing quality paper.

[0198] The enzymes, enzyme blends/compositions of the disclosure can be used to process or treat a number of other cellulosic material, including, e.g., fibers from wood, cotton, hemp, flax or linen.

[0199] Accordingly, the disclosure provides wood, wood pulp, paper, paper pulp, paper waste or wood or paper recycling treatment processes using an enzyme, enzyme blend/composition of the disclosure.

[0200] The enzymes, enzyme blends/compositions of the disclosure can be used for deinking printed wastepaper, such as newspaper, or for deinking noncontact-printed wastepaper, e.g., xerographic and laser-printed paper, and mixtures of contact and noncontact-printed wastepaper, as described in U.S. Patent No. 6,767,728 or 6,426,200; Neo, J. Wood Chem. Tech. 1986, 6(2):147. They can also be used to produce xylose from a paper-grade hardwood pulp in a process involving extracting xylan contained in pulp into a liquid phase, subjecting the xylan contained in the obtained liquid phase to conditions sufficient to hydrolyze xylan to xylose, and recovering the xylose. The extracting step, for example, can include at least one treatment of an aqueous suspension of pulp or an alkali-soluble material by an enzyme or an enzyme blend/composition (see, U.S. Patent No. 6,512,110). The enzymes, enzyme blends/compositions of the disclosure can be used to dissolve pulp from cellulosic fibers such as recycled paper products made from hardwood fiber, a mixture of hardwood fiber and softwood fiber, waste paper, e.g., from unprinted envelopes, de-inked envelopes, unprinted

ledger paper, de-inked ledger paper, and the like, as described in, *e.g.*, U.S. Patent No. 6,254,722.

6.4.2 Treating Fibers and Textiles

[0201] The disclosure provides methods of treating fibers and fabrics using one or more enzymes, enzyme blends/compositions of the disclosure. The enzymes, enzyme blends/compositions can be used in any fiber- or fabric-treating method, which are known in the art. See, *e.g.*, U.S. Patent Nos. 6,261,828; 6,077,316; 6,024,766; 6,021,536; 6,017,751; 5,980,581; U.S. Patent Publication No. 20020142438 A1. For example, enzymes, enzyme blends/compositions of the disclosure can be used in fiber and/or fabric desizing. The feel and appearance of a fabric can be, for example, improved by a method comprising contacting the fabric with an enzyme or enzyme blend/composition of the disclosure in a solution. Optionally, the fabric is treated with the solution under pressure. The enzymes, enzyme blends/composition of the disclosure can also be used to remove stains.

[0202] The enzymes, enzyme blends/compositions of the disclosure can be used to treat a number of other cellulosic material, including fibers (*e.g.*, fibers from cotton, hemp, flax or linen), sewn and unsewn fabrics, *e.g.*, knits, wovens, denims, yarns, and toweling, made from cotton, cotton blends or natural or manmade cellulose or blends thereof. The textile treating processes can be used in conjunction with other textile treatments, *e.g.*, scouring and/or bleaching. Scouring, for example, is the removal of non-cellulosic material from the cotton fiber, *e.g.*, the cuticle (mainly consisting of waxes) and primary cell wall (mainly consisting of pectin, protein and xyloglucan).

6.4.3 Treating Foods and Food Processing

[0203] The enzymes, enzyme blends/compositions of the disclosure have numerous applications in food processing industry. They can, for example, be used to improve extraction of oil from oil-rich plant material, *e.g.*, oil-rich seeds. The enzymes, enzyme blends/compositions of the disclosure can be used to extract soybean oil from soybeans, olive oil from olives, rapeseed oil from rapeseed, or sunflower oil from sunflower seeds.

[0204] The enzymes, enzyme blends/compositions of the disclosure can also be used to separate components of plant cell materials. For example, they can be used to separate plant cells into components. The enzymes, enzyme blends/ compositions of the disclosure can also be used to separate crops into protein, oil, and hull fractions. The separation process can be performed using known methods.

[0205] The enzymes, enzyme blends/compositions of the disclosure can, in addition to the uses above, be used to increase yield in the preparation of fruit or vegetable juices, syrups,

extracts and the like. They can also be used in the enzymatic treatment of various plant cell wall-derived materials or waste materials from, e.g., cereals, grains, wine or juice production, or agricultural residues such as, e.g., vegetable hulls, bean hulls, sugar beet pulp, olive pulp, potato pulp, and the like. Further, they can be used to modify the consistency and/or appearance of processed fruits or vegetables. Additionally, they can be used to treat plant material so as to facilitate processing of the plant material (including foods), purification or extraction of plant components. The enzymes, enzyme blends/compositions of the disclosure can be used to improve feed value, decrease the water binding capacity, improve the degradability in waste water plants and/or improve the conversion of plant material to ensilage, and the like.

[0206] The enzymes, enzyme blends/compositions of the disclosure can also be used in baking applications. In some embodiments, they are used to create non-sticky doughs that are not difficult to machines and to reduce biscuit sizes. They can also be used to hydrolyze arabinoxylans to prevent rapid rehydration of the baked product that can lead to loss of crispiness and reduced shelf-life. For example, they are used as additives in dough processing.

6.4.4 Animal Feeds and Food or Feed or Food Additives

[0207] The disclosure provides methods for treating animal feeds and foods and food or feed additives (supplements) using enzymes, enzyme blends/compositions of the disclosure. Animals including mammals (e.g., humans), birds, fish, and the like. The disclosure provides animal feeds, foods, and additives (supplements) comprising enzymes, enzyme blends/compositions of the disclosure. Treating animal feeds, foods and additives using enzymes of the disclosure can help in the availability of nutrients, e.g., starch, protein, and the like, in the animal feed or additive (supplements). By breaking down difficult to digest proteins or indirectly or directly unmasking starch (or other nutrients), the enzymes, enzyme blends/compositions can make nutrients more accessible to other endogenous or exogenous enzymes. They can also simply cause the release of readily digestible and easily absorbed nutrients and sugars.

[0208] When added to animal feed, enzymes, enzyme blends/compositions of the disclosure improve the *in vivo* break-down of plant cell wall material partly by reducing the intestinal viscosity (see, e.g., Bedford et al., Proceedings of the 1st Symposium on Enzymes in Animal Nutrition, 1993, pp. 73-77), whereby a better utilization of the plant nutrients by the animal is achieved. Thus, by using enzymes, enzyme blends/compositions of the disclosure in feeds, the growth rate and/or feed conversion ratio (*i.e.*, the weight of ingested feed relative to weight gain) of the animal can be improved.

[0209] The animal feed additive of the disclosure may be a granulated enzyme product which can be readily mixed with feed components. Alternatively, feed additives of the disclosure can form a component of a pre-mix. The granulated enzyme product of the disclosure may be

coated or uncoated. The particle size of the enzyme granulates can be compatible with that of the feed and/or the pre-mix components. This provides a safe and convenient mean of incorporating enzymes into feeds. Alternatively, the animal feed additive of the disclosure can be a stabilized liquid composition. This may be an aqueous- or oil-based slurry. See, e.g., U.S. Patent No. 6,245,546.

[0210] An enzyme, enzyme blend/composition of the disclosure can be supplied by expressing the enzymes directly in transgenic feed crops (e.g., as transgenic plants, seeds and the like), such as grains, cereals, corn, soy bean, rape seed, lupin and the like. As discussed above, the disclosure provides transgenic plants, plant parts and plant cells comprising a nucleic acid sequence encoding a polypeptide of the disclosure. The nucleic acid is expressed such that the enzyme of the disclosure is produced in recoverable quantities. The xylanase can be recovered from any plant or plant part. Alternatively, the plant or plant part containing the recombinant polypeptide can be used as such for improving the quality of a food or feed, e.g., improving nutritional value, palatability, and rheological properties, or to destroy an antinutritive factor.

[0211] The disclosure provides methods for removing oligosaccharides from feed prior to consumption by an animal subject using an enzyme, enzyme blend/composition of the disclosure. In this process a feed is formed to have an increased metabolizable energy value. In addition to enzymes, enzyme blends/compositions of the disclosure, galactosidases, cellulases, and combinations thereof can be used.

[0212] The disclosure provides methods for utilizing an enzyme, an enzyme blend/composition of the disclosure as a nutritional supplement in the diets of animals by preparing a nutritional supplement containing a recombinant enzyme of the disclosure, and administering the nutritional supplement to an animal to increase the utilization of hemicellulase contained in food ingested by the animal.

6.4.5 Waste Treatment

[0213] The enzymes, enzyme blends/compositions of the disclosure can be used in a variety of other industrial applications, e.g., in waste treatment. For example, in one aspect, the disclosure provides a solid waste digestion process using the enzymes, enzyme blends/compositions of the disclosure. The methods can comprise reducing the mass and volume of substantially untreated solid waste. Solid waste can be treated with an enzymatic digestive process in the presence of an enzymatic solution (including the enzymes, enzyme blends/compositions of the disclosure) at a controlled temperature. This results in a reaction without appreciable bacterial fermentation from added microorganisms. The solid waste is converted into a liquefied waste and any residual solid waste. The resulting liquefied waste can be separated from said any residual solidified waste. See, e.g., U.S. Patent No. 5,709,796.

6.4.6 Detergent, Disinfectant and Cleaning Compositions

[0214] The disclosure provides detergent, disinfectant or cleanser (cleaning or cleansing) compositions comprising one or more enzymes, enzyme blends/compositions of the disclosure, and methods of making and using these compositions. The disclosure incorporates all known methods of making and using detergent, disinfectant or cleanser compositions. See, e.g., U.S. Patent Nos. 6,413,928; 6,399,561; 6,365,561; 6,380,147.

[0215] In specific embodiments, the detergent, disinfectant or cleanser compositions can be a one- and two-part aqueous composition, a non-aqueous liquid composition, a cast solid, a granular form, a particulate form, a compressed tablet, a gel and/or a paste and a slurry form. The enzymes, enzyme blends/compositions of the disclosure can also be used as a detergent, disinfectant, or cleanser additive product in a solid or a liquid form. Such additive products are intended to supplement or boost the performance of conventional detergent compositions, and can be added at any stage of the cleaning process.

[0216] The present disclosure provides cleaning compositions including detergent compositions for cleaning hard surfaces, detergent compositions for cleaning fabrics, dishwashing compositions, oral cleaning compositions, denture cleaning compositions, and contact lens cleaning solutions.

[0217] When the enzymes of the disclosure are components of compositions suitable for use in a laundry machine washing method, the compositions can comprise, in addition to an enzyme, enzyme blend/composition of the disclosure, both a surfactant and a builder compound. They can additionally comprise one or more detergent components, e.g., organic polymeric compounds, bleaching agents, additional enzymes, suds suppressors, dispersants, lime-soap dispersants, soil suspension and antiredeposition agents, and corrosion inhibitors.

[0218] Laundry compositions of the disclosure can also contain softening agents, as additional detergent components. Such compositions containing carbohydrase can provide fabric cleaning, stain removal, whiteness maintenance, softening, color appearance, dye transfer inhibition and sanitization when formulated as laundry detergent compositions.

[0219] The disclosure thus further provides a process of saccharification a biomass material comprising hemicellulose. Such a biomass material can optionally comprise cellulose. Exemplary biomass materials include, without limitation, corcob, switchgrass, sorghum, and/or bagasse. Accordingly the disclosure provides a process of saccharification, comprising treating a biomass material herein comprising hemicellulose and optionally cellulose with an enzyme blend/composition as described herein. The enzyme blend/composition used in such a process include 0.5 g to 40 g (e.g., 0.5 g to 20 g, 0.5 g to 30 g, 0.5 g to 40 g, 0.5 g to 15 g, 0.5 g to 10 g, 0.5 g to 5 g, 0.5 g to 7 g, etc) of polypeptides having xylanase activity per kg of hemicellulose in the biomass material. The enzyme blend/composition used in such a process can also include 1 g to 40 g (e.g., 2 g to 20 g, 3 g to 7 g, 1 g to 5 g, or 2 g to 5 g, etc.) of polypeptides having xylanase activity per kg of hemicellulose in the biomass material. The

enzyme blend/composition used in such a process can include 0.5 g to 50 g (e.g., 0.5 g to 50 g, 0.5 g to 45 g, 0.5 g to 40 g, 0.5 g to 30 g, 0.5 g to 25 g, 0.5 g to 20 g, 0.5 g to 15 g, 0.5 g to 10 g, etc) of polypeptides having β -xylosidase activity per kg of hemicellulose in the biomass material. The enzyme blend/composition used in such a process can also include 1 g to 50 g (e.g., 2 g to 40 g, 4 g to 20 g, 4 g to 10 g, 2 g to 10 g, 3 g to 7 g, etc.) of polypeptide having β -xylosidase activity per kg of hemicellulose in the biomass material. The enzyme blend/composition used in such a process can include 0.2 g to 20 g (e.g., 0.2 g to 18 g, 0.2 g to 15 g, 0.3 g to 10 g, 0.2 g to 8 g, 0.2 g to 5 g, etc) of polypeptides having L- α -arabinofuranosidase activity per kg of hemicellulose in the biomass material. The enzyme blend/composition used in such a process can include 0.5 g to 20 g (e.g., 1 g to 10 g, 1 g to 5 g, 2 g to 6 g, 0.5 g to 4 g, or 1 g to 3 g, etc) of polypeptides having L- α -arabinofuranosidase activity per kg of hemicellulose in the biomass material. The enzyme blend/composition can also include 1 g to 100 g (e.g., 1 g to 100 g, 2 g to 80 g, 3 g to 50 g, 5 g to 40 g, 2 g to 20 g, 10 g to 30 g, or 12 g to 18 g, etc) of polypeptides having cellulase activity per kg of cellulose in the biomass material. Optionally, the amount of polypeptides having β -glucosidase activity can constitute up to 50% of the total weight of polypeptides having cellulase activity.

[0220] A suitable process preferably yields 60% to 90% xylose from the hemicellulose xylan of the biomass material treated. Suitable biomass materials include one or more of, for example, corncob, switchgrass, sorghum, and/or bagasse. As such, a process preferably yields at least 70% (e.g. at least 75%, at least 80%) xylose from hemicellulose xylan from one or more of these biomass materials. For example, the process yields 60% to 90% of xylose from hemicellulose xylan of a biomass material comprising hemicellulose, including, without limitation, corncob, switchgrass, sorghum, and/or bagasse.

[0221] The process optionally further comprises recovering monosaccharides.

[0222] The disclosure 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 disclosure in any way.

7. EXAMPLE 1: REPRESENTATIVE EXPERIMENTAL METHODS

7.1 Materials and Methods

[0223] The following assays/methods were used in Example 1 and subsequent examples. Any deviations from the protocols provided below are indicated.

7.1.1 Preparation of hemicellulose from plant tissues

[0224] Hemicellulose preparations were prepared using a modification of the NaOH/sonication procedure described by Erbringerova et al. (Carbohydrate Polymers 1998, 37:231). Dry plant material was ground to pass a 1 mm screen and 10 g of this material was suspended in 250 mL of 5% (wt/v) NaOH. The suspension was heated to 80°C without stirring for 30 min then sonicated for 15 min at ambient temperature using a probe sonicator at high setting. The suspension was returned to 80°C for an additional 30 min then allowed to cool to room temperature. Solids were removed from the suspension by centrifugation at 3000 x g for 15 min and the resulting supernatant was decanted into 1 L of ethanol which was cooled on ice. After 30 min the resulting precipitate was recovered by first decanting the clear liquid above the precipitate then filtration without allowing the precipitate to fully air dry. The filter cake was first washed with 200 mL of cold, 80% ethanol then removed from the filter without allowing it to air dry. The filter cake was re-dissolved in 200 mL of water and the pH of the solution was adjusted to 5.5 with acetic acid. The extracted carbohydrate was re-precipitated by addition to 1 L of ethanol on ice and the resulting precipitate was again recovered by filtration as above. The filter cake was frozen and remaining solvent and water was removed by lyophilization. Yield of recovered carbohydrate ranged from 6 to 23% of the starting plant material depending on the tissue and the preparation.

7.1.2 Dilute ammonia pretreatment of biomass substrates

[0225] Corncob and switchgrass were pretreated prior to enzymatic hydrolysis according to the methods and processing ranges described in WO06110901A (unless otherwise noted).

7.1.3 Compositional analysis of biomass

[0226] The 2-step acid hydrolysis methods described in "Determination of structural carbohydrates and lignin in the biomass" (National Renewable Energy Laboratory, Golden, CO 2008, available at <http://www.nrel.gov/biomass/pdfs/42618.pdf>) were used to measure the composition of biomass substrates. Enzymatic hydrolysis results are reported in terms of percent conversion with respect to the theoretical yield from the starting glucan and xylan content of the substrate.

7.1.4 Preparation of crude oligomers from ammonia pretreated corncob

[0227] Crude oligomers for screening hemicellulases were prepared from corncob by the following procedure. Hammer-milled corncob (-1/4 in mean diameter) plus 6% ammonia (w/w) was heated to 145°C with direct injection of steam into a stirred pressure reactor. After 20 min excess ammonia was flashed out of the reactor at a final vacuum of -0.1 bar. The ammonia pretreated cob was then placed in a sterile stirred reactor for enzyme saccharification. Enough water was added to obtain a final total solids loading of 25% (w/w) after all additions are made.

The pH of the reactor was maintained at pH 5.3 with 4 N sulfuric acid and the temperature controlled at 47°C. Spezyme® CP, Multifect® Xylanase (Danisco US Inc., Genencor), and Novo 188 (Novozymes, Denmark) were added at loadings of 20, 10 and 5 mg/g of cellulose, respectively, and allowed to saccharify the pretreated cob to sugars and oligomers for 116 h. The material was then cooled to 33°C and the pH adjusted to 5.8 with 4 N NaOH. The glucose and xylose were then fermented to ethanol by adding a seed culture of a recombinant *Zymomonas mobilis* strain (10% total volume, ATCC accession no. PTA-1798) as described in U.S. 7,354,755. The fermentation progress was followed until all of the glucose and -95% of the xylose was consumed. A 0.5 L aliquot of the fermentation broth was clarified by centrifugation (21,000 x g) for 20 min followed by filtration of the supernatant through a 0.2 micron filtering unit (Nalgene). The ethanol was removed from the filtered fermentation broth on a rotovap maintained at 35°C under house vacuum. The total volume of the final liquor was reduced by ~4x by the latter procedure.

7.1.5 Total protein assays

[0228] Different total protein determination methods were employed depending on the nature of the protein sample (*i.e.*, purified, fermentation broth, commercial product, etc.). The BCA protein assay is an example of a colorimetric assay that measures protein concentration with a spectrophotometer.

[0229] Reagents: BCA Protein Assay Kit (Pierce Chemical, Product #23227), 50 mM Sodium Acetate buffer pH 5.0, 15% trichloroacetic acid (TCA), 0.1 N NaOH, BSA stock solution, Reagent A, Reagent B (from protein assay kit)

[0230] Procedure: Enzyme dilutions were prepared in test tubes using 50 mM Sodium Acetate buffer. Diluted enzyme solution (0.1 mL) was added to 2 mL Eppendorf centrifuge tubes containing 1 mL 15% TCA. The tubes were vortexed and placed in an ice bath for 10 min. The samples were then centrifuged at 14,000 rpm for 6 min. The supernatant was poured out, the pellet resuspended in 1 mL 0.1 N NaOH, and the tubes vortexed until the pellet dissolved. BSA standard solutions were prepared from a stock solution of 2 mg/mL. BCA working solution was prepared by mixing 0.5 mL Reagent B with 25 mL Reagent A. The resuspended protein (0.1 mL each) was added to 3 Eppendorf centrifuge tubes. Two mL Pierce BCA working solution was added to each of the sample and serially diluted BSA standard Eppendorf tubes. All tubes were incubated in a 37°C water bath for 30 min. The samples were then cooled to room temperature (15 min) and the absorbance measured at 562 nm in a spectrophotometer.

[0231] Calculations: Average values for each BSA protein standard absorbance were calculated and plotted, absorbance on x-axis and concentration (mg/mL) on the y-axis. A linear curve fit was applied and the equation for the line calculated using the formula: $y=mx+b$

[0232] The raw concentration of the enzyme samples was calculated by substituting the absorbance for the x- value. The total protein concentration was calculated by multiplying with

the dilution factor.

[0233] The total protein of purified samples was determined by A280 (see, e.g., Pace et al., Protein Science, 1995, 4:2411).

[0234] Some protein samples were measured using the Biuret method as modified by Weichselbaum and Gornall using Bovine Serum Albumin as a calibrator (modified Biuret) (Weichselbaum, Amer. J. Clin. Path. 1960, 16:40; Gornall et al., J. Biol. Chem. 1949, 177:752).

[0235] Total protein content of fermentation products was also sometimes measured as total nitrogen by combustion, capture and measurement of released nitrogen, either by Kjeldahl (rtech laboratories, www.rtechlabs.com) or in-house by the DUMAS method (TruSpec CN, www.leco.com) (SADER, et al. Archives of Veterinary Science, 9(2):73-79, 2004). For complex protein-containing samples, e.g. fermentation broths, an average 16% N content, and the conversion factor of 6.25 for nitrogen to protein was used. In some cases, total precipitable protein was measured to remove interfering non-protein nitrogen. A 12.5 % final TCA concentration was used and the protein-containing TCA pellet was resuspended in 0.1 M NaOH.

[0236] In other cases, Coomassie Plus- the Better Bradford Assay (Thermo Scientific, Rockford, IL product #23238) was used according to manufacturer recommendation.

7.1.6 Synthetic substrate (para-nitrophenyl substrate) activity assays

[0237] Active protein from *T. reesei* expression of cloned genes was confirmed with model substrate assays. Cellulase and hemicellulase activities on the synthetic substrates, such as 4-nitrophenyl α -L-arabinofuranoside (pNPA, Sigma N3641) and 4-nitrophenyl β -D-glucopyranoside (pNPG, Sigma N7006), and 4-nitrophenyl β -D-xylopyranoside (pNPX, Sigma N2132) were measured as follows: Substrate solution was prepared by dissolving 30 mg of synthetic substrate in 100 mL 50 mM Sodium Acetate buffer, pH 4.8. Sodium carbonate (1 M) was prepared for reaction quenching. Substrate solution (100 μ L) was dispensed into Costar 96 well plates (Cat no. 9017). 20 μ L of enzyme sample was dispensed into a microtiter plate well. The microtiter plate was incubated at 50°C for 10 min using a Thermomixer R heating and cooling shaker (Eppendorf). 50 μ L of 1 M sodium carbonate was added to each well to quench the reaction. Absorbance at 400 nm wavelength was read with SpectraMax 340C384 Microplate Spectrophotometer (Molecular Devices). Units per mL were determined by using a *p*-nitrophenol standard curve. A Quad-delete *Trichoderma* host, from which the *cbh1*, *cbh2*, *egl1* and *egl2* genes were deleted (see WO 05/001036), was analyzed with the enzyme samples as a control for the activity of enzymes expressed in this background.

7.1.7 Cob saccharification assay

[0238] Typically, Corncob saccharification in a microtiter plate format was performed in accordance with the following procedures. The biomass substrate, dilute ammonia pretreated corn cob, was diluted in water and pH-adjusted with sulfuric acid to create a pH 5, 7% cellulose slurry that was used directly in the assay. The enzymes tested included: commercial cellulase products, e.g. Accellerase® 1000, Accellerase® 1500 (Danisco US Inc., Genencor), *T. reesei* fermentation broths, and purified enzymes. The enzymes were loaded based on mg total protein per gram of cellulose (as determined by compositional analysis) in the corn cob substrate. The enzymes were diluted in 50 mM Sodium Acetate pH 5.0 to obtain the desired loading concentration at the required volume. Forty microliters of enzyme solution was added to 70 mg of dilute-ammonia pretreated corn cob at 7% cellulose per well (equivalent to 4.5% cellulose final). The assay plate was incubated at room temperature for 10 min. The assay plates were covered with aluminum plate sealers and the plates incubated at 50°C, 200 rpm, for three days. At the end of the incubation period, the saccharification reaction was quenched by adding 100 µL of 100 mM glycine buffer, pH 10.0 per well and the plate was centrifuged for 5 min at 3,000 rpm. Ten microliters of the supernatant were added to 200 µL of MilliQ water in a 96-well HPLC plate and the soluble sugars were measured by HPLC.

[0239] This describes a typical method that was used in multiple Examples herein. In certain Examples, corn cob saccharification was measured using a modified protocol. The modifications are described with the individual examples.

7.1.8 Sugar analysis by HPLC

[0240] Typically, samples from cob saccharification hydrolysis were prepared by centrifugation to clear insoluble material, filtration through a 0.22 µm nylon filter (Spin-X centrifuge tube filter, Corning Incorporated, Corning, NY) and dilution to an appropriate concentration of soluble sugars with distilled water. Monomer sugars were determined on a Shodex Sugar SH-G SH1011, 8x300mm with a 6x50 mm SH-1011P guard column (www.shodex.net). Solvent was 0.01 N H₂SO₄ run at 0.6 mL/min. Column temperature was 50°C and detection was made using a refractive index detector. External standards of glucose, xylose and arabinose were run with each sample set. Certain examples herein use a protocol to achieve the same end with a somewhat modified set of protocols. The specific modifications to the protocols are described with individual examples.

[0241] Oligomeric sugars were separated by size exclusion chromatography using a Tosoh Biosep G2000PW column 7.5 mm x 60 cm (www.tosohbioscience.de). The solvent was distilled water at 0.6 mL/min and the column was run at room temperature. Six carbon sugar standards used for size calibration were: stachyose, raffinose, cellobiose and glucose; and 5 carbon sugars were: xylohexose, xylopentose, xylotetrose, xylotriose, xylobiose and xylose. Xylo-oligomers were obtained from Megazyme (www.megazyme.com). Detection was by refractive index and when reported quantitatively results are either as peak area units or relative peak areas by percent.

[0242] Total soluble sugars were determined by acid hydrolysis of the centrifuged and filter clarified samples described above. The clarified sample was diluted 1:1 with 0.8 N H₂SO₄ and the resulting solution was autoclaved in a capped vial for a total cycle time of 1 h at 121 °C. Results are reported without correction for loss of monomer sugar during the hydrolysis.

7.1.9 Protein Analysis by HPLC

[0243] To separate and quantify the enzymes contained in broth from 14L fermentations of the integrated expression strains, liquid chromatography (LC) and mass spectroscopy (MS) were performed. Enzyme samples were first treated with a recombinantly expressed endoH glycosidase from *S. plicatus* (e.g., NEB P0702L). EndoH was used at a ratio of 0.01-0.03 µg endoH protein per µg sample total protein and incubated for 3 h at 37°C, pH 4.5-6.0 to enzymatically remove N-linked glycosylation prior to HPLC analysis. Approximately 50 µg of protein was then injected for hydrophobic interaction chromatography using an Agilent 1100 HPLC system with an HIC-phenyl column and a high-to-low salt gradient over 35 min was performed on samples of concentrated fermentation broth. The gradient was achieved using high salt buffer A: 4 M ammonium sulphate containing 20 mM potassium phosphate pH 6.75 and low salt buffer B: 20 mM potassium phosphate pH 6.75. Peaks were detected with UV light at 222 nm and fractions were collected and identified by mass spectroscopy.

7.1.10 Cellulase Activity Assay Using Calcofluor White

[0244] Cellulase activity was measured on PASC using a calcofluor white detection method (Appl. Biochem. Biotechnol. 161:313-317). All chemicals used were of analytical grade. Avicel PH-101 was purchased from FMC BioPolymer (Philadelphia, PA). Calcofluor white was purchased from Sigma (St. Louis, MO). Phosphoric acid swollen cellulose (PASC) was prepared from Avicel PH-101 using an adapted protocol of Walseth, TAPPI 1971, 35:228 and Wood, Biochem. J. 1971, 121:353-362. In short, Avicel was solubilized in concentrated phosphoric acid then precipitated using cold deionized water. After the cellulose was collected and washed with more water to neutralize the pH, it was diluted to 1% solids in 50 mM Sodium Acetate buffer, pH 5.0.

[0245] All enzyme dilutions were made into 50 mM Sodium Acetate buffer, pH 5.0. GC220 Cellulase (Danisco US Inc., Genencor) was diluted to 2.5, 5, 10, and 15 mg protein/g PASC, to produce a linear calibration curve. Samples to be tested were diluted to fall within the range of the calibration curve, i.e. to obtain a response of 0.1 to 0.4 fraction product. 150 µL of cold 1% PASC was added to 20 µL of enzyme solution in 96-well microtiter plates. The plate was covered and incubated for 2 h at 50°C, 200 rpm in an Innova incubator/shaker. The reaction was quenched with 100 µL of 50 µg/mL Calcofluor in 100 mM Glycine, pH 10. Fluorescence was read on a fluorescence microplate reader (SpectraMax M5 by Molecular Devices) at excitation wavelength Ex = 365 nm and emission wavelength Em = 435 nm. The result (shown

in Figure 25) is expressed as the fraction product according to the equation:

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

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

7.1.11 Cultivation of *Fusarium verticillioides* and purification of major hemicellulase activities detected in the extracellular protein

[0246] Wild type *Fusarium verticillioides* source was as described in Table 1 of Fuchs et al., Fungal Genetics and Biology 2004, 41:852-863. The fungus was grown using destarched corn grain fiber using a modification of the method described in Li et al., Applied Biochemistry and Biotechnology 2005, (121-124):321-334.

[0247] Starch was removed from the dried corn pericarp fraction from a corn dry mill fractionation by slurrying 200 g of corn pericarp in 3 L of tap water and heating to 80 °C with 5 mL (500 mg) of high temperature amylase Liquozyme SC DS (Novozymes, Denmark). The mixture was stirred occasionally with a spatula and held at 80°C for 30 min then allowed to cool to room temperature for about 2 h. The resulting slurry was decanted onto a 20 mesh screen to partially de-water. The solids on the screen were further washed with 4 L of tap water then air dried overnight before a final drying in an oven at 60°C. The dried material was ground to pass a 1 mm screen in a knife mill before use.

[0248] The *F. verticillioides* culture was maintained on potato dextrose agar (Sigma P6685) and a rich growth media of 24 g/L potato dextrose was inoculated with mycelia from the plate. The growth culture was incubated for 3 days at 30°C with agitation at 130 rpm. After 3 days growth the 150 mL of the resulting cell mass and spent media were used to inoculate a corn pericarp induction media. The induction media was 80 g destarched corn pericarp in 1 L base media (15 g KH₂PO₄, 5 g ammonium sulfate, 20 g yeast extract, 0.5 g magnesium sulfate and 1 g Tween 80 at pH4.8). The culture was maintained at 30°C with agitation at 130 rpm for 7 days.

[0249] Extracellular protein was separated from the spent grain and fungal cell mass by centrifugation at 2,000 x G for 20 min. The partially clarified supernatant was passed first through a 0.45 µm filter and then through a 0.22 µm filter. The clarified filtrate was stored at 4°C until used. The filtrate was concentrated approximately 14-fold with Amicon ultra-4 centrifugal filters with a 30 kD cutoff (Millipore).

[0250] L-α-arabinofuranosidase (LARF) and β-xylosidase (BXL) activities were purified from the fungal supernatant by first exchanging the proteins into 50 mM pH 6.3 (N-Morpholino) ethanesulphonic acid-NaOH (MES) buffer, then applying protein onto a GE Health Sciences Q-Sepharose Fast Flow quaternary amine anion exchange column (20 mL bed volume equilibrated with MES buffer). The protein purification system used was a GE Health Sciences Akta system using Unicorn software. Proteins were eluted with a 0-0.5 M NaCl gradient in MES buffer.

[0251] Activity in column fractions was monitored using either 4-nitrophenyl β -D-xylopyranoside (for BXL activity) or 4-nitrophenyl α -L-arabinofuranoside (for LARF activity). In both cases, 2 μ L of the column fraction was added to 95 μ L of 50 mM pH 5.0 Sodium Acetate buffer and 5 μ L of the p-nitrophenyl substrate, mixed and incubated 2 min at 23°C. 100 μ L of 1 N sodium carbonate stop solution was added and absorbance at 405 nm was read.

[0252] Active enzyme fractions from the anion exchange system were collected, concentrated on Amicon Centriprep filtration, and applied to a GE Health Sciences SP-Sepharose Fast Flow sulfopropyl cation exchange column (20 mL bed equilibrated with Sodium Acetate buffer) using the same protein purification system described above. Proteins were eluted with a 0-1 M NaCl gradient in Sodium Acetate buffer, and activity in the fractions was monitored as described above. Active fractions were then collected, pooled and concentrated if necessary. Proteins in active fractions were identified by MS-MS.

7.1.12 Protein Identification by MS-MS

[0253] The identification of a representative set of *Fusarium verticillioides* family GH54 L- α -arabinofuranosidase (LARF), family GH51 LARF, two family GH3 β -xylosidases and a family GH30 β -xylosidase (BXL) is described below.

[0254] For the identification of GH54 LARF, 200 μ L of 1 N HCl (EM Industries, Hawthorne, NY) was added to 250 μ L of "Fraction B3" (0.35 mg protein) from the purification procedure described above and the sample iced for 10 min. Next, 200 μ L of 50% TCA (Sigma-Aldrich, St. Louis, MO) was added and samples were iced for another 10 min to precipitate protein. Samples were centrifuged for 2 min at 16,000 rcf and the supernatants were discarded. The pelleted protein was washed with 90% cold acetone (J.T. Baker, Phillipsburg, NJ), then centrifuged for 1 min at 16,000 rcf. The supernatant was discarded and the pellet was left to air dry. First, 30 μ L of 8 M urea (MP Biomedicals, Solon, OH) was added to the pellet, then 4 μ L of 0.2 M DTT (Sigma-Aldrich, St. Louis, MO). Sample was incubated at 52°C for 30 min and then 4 μ L of 0.44 M iodoacetamide (Sigma-Aldrich, St. Louis, MO) was added and the solution incubated at room temperature in the dark for 30 min. Next, 120 μ L of 0.1% n-octyl β -D-glucopyranoside water (Sigma-Aldrich, St. Louis, MO) was added slowly, and then sample was divided into two equal parts. To half the sample, 7.5 μ g Trypsin (Promega Corp., Madison, WI) was added, and to the other half, 4 μ g of AspN (Roche Applied Science, Indianapolis, IN) was added. Both samples were incubated at 37°C for 1 h. Samples were quenched with 10% Trifluoroacetic Acid (Thermoscientific, Waltham, MA) for a final volume of 0.1% TFA. Both samples were run on a Thermofinnigan (San Jose, CA) LCQ-Deca electrospray ionization (ESI) ion-trap mass spectrometer. A Vydac C18 column (5 μ m, 300A, 0.2 x 150 mm, Michrom Bioresources, Auburn, CA) was run at a flow rate of 200 μ L/min. The injection volumes were 50 μ L, and were filtered through an on-line trapping cartridge (Peptide CapTrap, Michrom Bioresources, Auburn, CA) before loading onto the column. Separation of the digests was performed with a gradient as follows (Solvent A: 0.1% trifluoroacetic acid in H₂O (J.T. Baker,

Phillipsburg, NJ), Solvent B: 0.08% trifluoroacetic acid in acetonitrile (J.T. Baker, Phillipsburg, NJ)). For the identification of GH51 LARF, the two GH3 β -xylosidases, and GH30 β -xylosidase, the method described above was used, except samples were digested with Chymotrypsin (7.5 μ g, Sigma-Aldrich, St. Louis, MO) in addition to Trypsin and AspN. The TurboSequest search engine within Bioworks 3.31 (Thermo, San Jose) was used to identify proteins from the *Fusarium verticillioides* protein database. The protein database was downloaded from the BROAD Institute at http://www.broadinstitute.org/annotation/genome/fusarium_verticillioides/MultiDownloads.html.

7.1.13 Induction of *Fusarium verticillioides* Enzymes by Dilute Aqueous Ammonia Pretreated Switchgrass

[0255] For induction of *Fusarium verticillioides* enzymes in response to dilute aqueous ammonia pretreated switchgrass, switchgrass was ground to <1 mm diameter, then pretreated with 6% ammonia based on dry weight at 50% initial dry matter at 160°C maximum temperature for 90 min. The resulting material was 43.3% solids and contained 0.336% residual ammonia.

[0256] Wild type *F. verticillioides* culture was maintained on potato dextrose agar (Sigma P6685) and a rich growth media of 24 g/L potato dextrose was inoculated with mycelia from the plate. The growth culture was incubated for 3 days at 24°C with agitation at 140 rpm, during which time it became turbid with *Fusarium* cells. After 3 days growth, 4 flasks each containing 100 mL of base Christakopoulos media (0.1% KH_2PO_4 , 0.03% CaCl_2 , 0.03% $\text{MgSO}_4 \times 7 \text{ H}_2\text{O}$, 2.61% $\text{Na}_2\text{HPO}_4 \times 7 \text{ H}_2\text{O}$, 0.134% $\text{NaH}_2\text{PO}_4 \times 1 \text{ H}_2\text{O}$, 1.0% anhydrous ammonium phosphate) were inoculated with 4 mL of the resulting cell suspension. To the suspension, two grams dry matter of dilute aqueous ammonia-pretreated switchgrass was added as the sole carbon source.

[0257] After addition of the pretreated switchgrass, the pH was adjusted to 6.5 and the flask was swirled at 180 rpm at 23°C for 168 h. Prior to size exclusion chromatography, samples were analyzed at different time points by Bradford assay for protein, p-nitrophenyl arabinofuranosidase, p-nitrophenyl xylosidase and p-nitrophenyl glucosidase activities, and by SDS-PAGE gels for induction of enzymes. Intact non-pretreated switchgrass, 2% glucose and no carbon source were included as parallel controls and were found to lead to much lower levels of enzyme induction than was the case for dilute aqueous ammonia-pretreated switchgrass carbon source.

7.1.14 Size exclusion chromatography fractionation of the *Fusarium verticillioides* culture broth

[0258] The *Fusarium verticillioides* culture media containing the expressed enzymes from the

168-h induction described above was centrifuged at 3,500 x g to remove debris and cells. The supernatant was filtered through a 0.4 µm filter yielding a clear deep yellow liquid. The liquid was concentrated (~4X) two times in a large Amicon ultra 10K MW cutoff concentrator (UFC901024). A 1.7 mL concentrated sample was set aside for size exclusion chromatography (SEC) and assayed for protein content by BCA assay as described earlier. The concentrated protein sample was loaded onto a Superdex 16/60 SEC column equilibrated with 50 mM Sodium Acetate buffer pH 5.0. The column was eluted at 1 mL/min at 4°C. Absorbance of the sample at 214 nm and 280 nm were monitored. After elution of the void volume, 1 mL fractions were collected in a deep well polypropylene microtiter plate. The protein-containing fractions were aggregated into a separate deep well plate and all fractions were measured for total protein concentration by BCA assay, and p-nitrophenyl arabinosidase, p-nitrophenyl xylosidase, and p-nitrophenyl glucosidase activities as described in Example 1 (Synthetic substrate activity assays).

7.1.15 Purification of A NUMBER OF GH43, GH51 and GH3 HOMOLOGS (SECTION 8.1 BELOW) from *T. reesei* fermentation by cation exchange chromatography

[0259] Shake flask-scale enzyme production from wild type *Trichoderma reesei* cultures was performed as described in WO 2005/001036. The extracellular enzyme preparations were concentrated by Amicon Centriprep filtration if necessary, and applied to a GE Health Sciences SP-Sepharose Fast Flow sulfopropyl cation exchange column (20 mL bed equilibrated with Sodium Acetate buffer) using a GE Health Sciences Akta system protein purification system using Unicorn software. Proteins were eluted with a 0-1 M NaCl gradient in Sodium Acetate buffer and activity in the fractions were monitored as described above. Active fractions were then collected, pooled and concentrated if necessary. Activity in column fractions was monitored using either 4-nitrophenyl β-D-xylopyranoside (for BXL activity) or 4-nitrophenyl α-L-arabinofuranoside (for LARF activity). In both cases, 2 µL of the column fraction was added to 95 µL of 50 mM pH 5.0 Sodium Acetate buffer and 5 µL of the p-nitrophenyl substrate, mixed and incubated 2 min at 23°C. One hundred microliters of 1 N sodium carbonate stop solution was added and the absorbance was measured at 405 nm. In some cases an anion exchange protein purification step was useful before the cation exchange. In this case, the enzyme fractions were exchanged into 50 mM pH 6.3 (N-Morpholino) ethanesulphonic acid-NaOH (MES) buffer, and then the protein sample was applied onto a GE Health Sciences Q-Sepharose Fast Flow quaternary amine anion exchange column (20 mL bed volume equilibrated with MES buffer). The Akta protein purification system as described for the cation exchange step was used. Proteins were eluted with a 0-0.5 M NaCl gradient in MES buffer. The active enzymes were then exchanged into 50 mM, pH 5.0 Sodium Acetate buffer before cation exchange chromatography.

7.1.16 Purification of Xylanases

[0260] Xylanase purification was conducted in two stages:

Stage 1: Hydrophobic interaction chromatography (HIC). $(\text{NH}_4)_2\text{SO}_4$ was added to fermentation supernatant to achieve a final concentration of 1 M. The separation column used was a HiPrep Phenyl (highsub), 16/10, 20 mL. *Buffer A*: 20 mM sodium phosphate, pH 6.0. *Buffer B*: Buffer A + 1 M $(\text{NH}_4)_2\text{SO}_4$. Stepwise Elution: 45% B; 0% B; water+10% glycerol.

Stage 2: Gel filtration (GF). Eluate of 0% B from the HIC column was collected and loaded on a HiLoad 26/60 Superdex 75 prep grade (320 mL, GE Healthcare) column. The mobile phase was 20 mM sodium phosphate, pH 6.8, containing 0.15 M NaCl. The elution profile is shown in Figure 26. Figure 27 shows the SDS-PAGE detection of the two step separation of AfuXyn 5.

[0261] The Applied Biosystems BIOCAD® Vision and Amersham Pharmacia Biotech AKTA Explorer were used for protein purification of *T. reesei* XYN3. Approximately 150 mL of Xylanase 3, from an ultrafiltration concentrate, was loaded onto a Sephadex G-25M desalting column (total volume ~525 mL) equilibrated in 10 mM TES buffer, pH 6.8. 300-400 mL of this desalted sample was then loaded onto an anion-exchange column (high density quaternary amine resin; Applied Biosystems Inc.). The bound protein was then eluted using a salt gradient between 0-1 M sodium chloride using 8-column volumes of 25 mM TES buffer. Elution of the protein occurred between 0-250 mM sodium chloride, and was detected using 10% Bis-Tris NUPAGE® SDS-PAGE (Novex).

[0262] The eluted Xylanase 3-containing samples were concentrated to 10 mL with Vivaspin 5,000 MWCO (molecular weight cut-off) membrane concentrators (Vivascience; GE Healthcare). The concentrate was then loaded onto a High Load 26/60 Superdex 200 (ID no 17-1071-01; GE Healthcare). The column was equilibrated with 25 mM TES buffer, pH 6.8, containing 100 mM sodium chloride, and the bound protein was eluted over 8-column volumes of the TES buffer with sodium chloride.

[0263] The purity of the eluted protein was assessed using 10% Bis-Tris NUPAGE® SDS-PAGE (Novex), and determined to be greater than 95% pure.

7.1.17 Purification of Fv43D

[0264] The ultrafiltration concentrate (UFC) of *Fusarium verticillioides* 43D was buffer exchanged and dialyzed against 50 mM Sodium Acetate buffer, pH 4.0, overnight. The dialyzed material was passed through a HiTrap 1 mL column prepacked with sulfopropyl sepharose fast flow resin (GE Healthcare) designed for use with a syringe. The purified Fv43D was eluted with 50 mM Sodium Acetate buffer, pH 4.0, with 250 mM sodium chloride. The UFC and Sodium Acetate buffer were pushed through the column using a 5 mL syringe fitted with a GE Healthcare connector. The purified Fv43D was dialyzed overnight against 50 mM Sodium

Acetate buffer, pH 4.0. The purified protein was assayed by SDS-PAGE, HPLC, and mass spectroscopy to demonstrate homogeneity.

7.1.18 Purification of Fv51A

[0265] The ultrafiltration concentrate (UFC) of *Fusarium verticillioides* 51A was buffer exchanged and dialyzed against 50 mM Sodium Acetate buffer, pH 5.0, overnight. The dialyzed material was passed through a RESOURCE 15 6 mL column prepacked with methyl sulfonate media (GE Healthcare). The UFC was loaded at 1 mL/min against 50 mM Sodium Acetate buffer, pH 5.0, and eluted at 5 mL/min against 50 mM Sodium Acetate buffer, pH 5.0, using a 0 to 250 mM sodium chloride gradient. The eluted fractions were collected and assayed by SDS-PAGE. Fractions with purified Fv51A were concentrated using a 10,000 MWCO Vivaspin concentrator from Sartorius Stedim Biotech. The purified Fv51A was dialyzed against 50 mM Sodium Acetate buffer, pH 5.0, overnight. The purified protein was assayed by SDS-PAGE, HPLC, and mass spectroscopy to demonstrate homogeneity. The AKTA Explorer 100 system from GE Healthcare was used for the purification of Fv51A.

8. EXAMPLE 2: EXPRESSION OF INDIVIDUAL HEMICELLULASE GENES FROM VARIOUS SPECIES IN *TRICHODERMA REESEI*

8.1 *Fusarium verticillioides* genes

[0266] The sequence for Fv51A was obtained by searching the *Fusarium verticillioides* genome in the Broad Institute database (<http://www.broadinstitute.org/>) for GH51 arabinofuranosidase homologs.

[0267] The following genes from *Fusarium verticillioides* were expressed in *Trichoderma reesei*: Fv3A, Fv43A, Fv43B, Fv43D, Fv51A, Fv3B, Fv43C, Fv39A, Fv43E, Fv30A, Fv30B, and Fv43F. Fv3A sequence was obtained by searching for GH3 β -xylosidase homologs in *Fusarium verticillioides* genome. The annotated sequence lacked a signal sequence and the gene prediction program Augustus (<http://augustus.gobics.de/>) was used to identify upstream sequence which contained a signal sequence. Sequences for Fv39A, Fv43A, Fv43B, Fv43D, Fv43E, and Fv30A were obtained by searching the *Fusarium verticillioides* genome for GH39, GH30, and GH43 β -xylosidase homologs.

[0268] Open reading frames of the hemicellulase genes of interest were amplified by PCR using purified/extracted genomic DNA from *Fusarium verticillioides* as the template. The PCR thermocycler used was DNA Engine Tetrad 2 Peltier Thermal Cycler (BioRad Laboratories). The DNA polymerase used was PfuUltra II Fusion HS DNA Polymerase (Stratagene). The primers used to amplify the open reading frames were as follows:

Fv3A:

Forward primer MH124 (5'- CACCCATGCTGCTCAATCTTCAG-3') (SEQ ID NO:52)

Reverse primer MH125 (5'-TTACGCAGACTTGGGGTCTTGAG-3') (SEQ ID NO:53)

Fv43A:

Forward primer MH075 (5'-CACCATGTGGCTGACCTCCCCATT-3') (SEQ ID NO:54)

Reverse primer MH076 (5'-TTAGCTAAACTGCCACCAGTTGAAGTTG -3') (SEQ ID NO:55)

Fv43B:

Forward primer MH077 (5'-CACCATGCGCTTCTCTTGGCTATTGT-3') (SEQ ID NO:56)

Reverse primer MH078 (5'-CTACAATTCTGATTTCACAAAAACACC-3') (SEQ ID NO:57)

Fv43D:

Forward primer MH081 (5'-CACCATGCAGCTCAAGTTTCTG-3') (SEQ ID NO:58)

Reverse primer MH082 (5'-CTAAATCTTAGGACGAGTAAGC-3') (SEQ ID NO:59)

Fv51A:

Forward primer SK1159 (5'-CACCATGGTTCGCTTCAGTTCAATCCTAG-3') (SEQ ID NO:60)

Reverse primer SK1160 (5'-CTAGCTAGAGTAAGGCTTTCC-3') (SEQ ID NO:61)

Fv39A:

Forward: MH116 (5'-CACCATGCACTACGCTACCCTCACCAC-3') (SEQ ID NO:62)

Reverse: MH117 (5'-TCAAGTAGAGGGGCTGCTCACC-3') (SEQ ID NO:63)

Fv3B:

Forward primer MH126 (5'- CAC CAT GAA ACT CTC TAG CTA CCT CTG -3') (SEQ ID NO:64)

Reverse primer MH127 (5'- CTA CGA AAC TGT GAC AGT CAC GTT G -3') (SEQ ID NO:65)

Fv30A:

Forward primer MH112 (5'-CAC CAT GCT CTT CTC GCT CGT TCT TCC TAC-3') (SEQ ID NO:66)

Reverse primer MH113 (5'-TTA GTT GGT GCA GTG GCC ACG-3') (SEQ ID NO:67)

Fv30B:

Forward primer MH114 (5'-CAC CAT GAA TCC TTT ATC TCT CGG CCT TG-3') (SEQ ID NO:68)

Reverse primer MH115 (5'-CAG CCC TCA TAG TCG TCT TCT TC-3') (SEQ ID NO:69)

Fv43C:

Forward primer MH079 (5'- CAC CAT GCG TCT TCT ATC GTT TCC -3') (SEQ ID NO:70)

Reverse primer MH080 (5'- CTA CAA AGG CCT AGG ATC AA -3') (SEQ ID NO:71)

Fv39A:

Forward primer MH116 (5'-CAC CAT GCA CTA CGC TAC CCT CAC CAC-3') (SEQ ID NO:72)

Reverse primer MH117 (5'-TCA AGT AGA GGG GCT GCT CAC C-3') (SEQ ID NO:73)

Fv43E:

Forward primer MH147 (5'-CAC CAT GAA GGT ATA CTG GCT CGT GG-3') (SEQ ID NO:74)

Reverse primer MH148 (5'-CTA TGC AGC TGT GAA AGA CTC AAC C-3') (SEQ ID NO:75)

Fv43F:

Forward primer MH149 (5'-CACCATGTGGAACTCCTCGTCAGC-3') (SEQ ID NO:76)

Reverse primer MH150 (5'-CTA ATA AGC AAC AGG CCA GCC ATT G-3') (SEQ ID NO:77)

[0269] The forward primers included four additional nucleotides (sequences - CACC) at the 5'-end to facilitate directional cloning into pENTR/D-TOPO (Invitrogen, Carlsbad, CA) (Figure 28). The PCR conditions for amplifying the open reading frames were as follows (except for Fv51A): Step 1: 94°C for 2 min. Step 2: 94°C for 30 sec. Step 3: 57°C for 30 sec. Step 4: 72°C for 30-45 sec. Steps 2, 3 and 4 were repeated for an additional 29 cycles. Step 5: 72°C for 2 min. For Fv51A, the following conditions were used: Step 1: 94°C for 2 min. Step 2: 94°C for 30 sec. Step 3: 56°C for 30 sec. Step 4: 72°C for 45 sec. Steps 2, 3, 4 were repeated for an additional 25 cycles. Step 5: 4°C hold.

[0270] The PCR products of the corresponding hemicellulase open reading frames were purified using a Qiaquick PCR Purification Kit (Qiagen, Valencia, CA). The purified PCR products were cloned into the pENTR/D-TOPO vector, transformed into TOP10 chemically competent *E. coli* cells (Invitrogen, Carlsbad, CA) and plated on LA plates with 50 ppm kanamycin. Plasmid DNA was obtained from the *E. coli* transformants using a QIAspin plasmid preparation kit (Qiagen). Sequence data for the DNA inserted in the pENTR/D-TOPO vector was obtained using M13 forward and reverse primers (Sequetech, Mountain View, CA). A pENTR/D-TOPO vector with the correct DNA sequence of the corresponding hemicellulase open reading frame was recombined with the pTrex3gM destination vector (WO 05/001036, Figure 29) using LR clonase reaction mixture (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. The product of the LR clonase reaction was subsequently

transformed into TOP10 chemically competent *E. coli* cells which were then plated on LA containing 50ppm carbenicillin. The resulting expression vectors were pTrex3gM plasmids containing the corresponding hemicellulase open reading frames that resulted from the recombination event between the attR1 and attR2 sites of pTrex3gM and the attL1 and attL2 sites of pENTR/D-TOPO; and the *Aspergillus nidulans* acetamidase selection marker (*amdS*). DNA of the expression vectors containing the corresponding hemicellulase open reading frames were isolated using a Qiagen miniprep kit and used for biolistic transformation of *Trichoderma reesei* spores.

[0271] Biolistic transformation of *Trichoderma reesei* with the pTrex3gM expression vector containing the corresponding hemicellulase open reading frame was performed using the following protocol. Transformation of the *Trichoderma reesei* cellulase quad delete ($\Delta cbh1$, $\Delta cbh2$, $\Delta eg1$, $\Delta eg2$) strain by helium-bombardment was accomplished using a Biolistic® PDS-1000/He Particle Delivery System from Bio-Rad (Hercules, CA) following the manufacturer's instructions (see patent publications WO 05/001036 and US 2006/0003408). Transformants were transferred to fresh acetamide selection plates (see patent publication WO 2009114380). Stable transformants were inoculated into filter microtiter plates (Corning), containing 200 μ L/well of Glycine Minimal media (6.0 g/L glycine; 4.7 g/L $(\text{NH}_4)_2\text{SO}_4$; 5.0 g/L KH_2PO_4 ; 1.0 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$; 33.0 g/L PIPPS; pH 5.5) with post sterile addition of -2% glucose/sophorose mixture (US 7,713,725) as the carbon source, 10 mL/L of 100 g/L of CaCl_2 , 2.5 mL/L of *T. reesei* trace elements (400X): 175 g/L Citric acid anhydrous; 200 g/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$; 16 g/L $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$; 3.2 g/L $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$; 1.4 g/L $\text{MnSO}_4 \cdot \text{H}_2\text{O}$; 0.8 g/L H_3BO_3). Transformants were grown in liquid culture for 5 days in an O_2 -rich chamber housed in a 28°C incubator. The supernatant samples from the filter microtiter plate were collected on a vacuum manifold. Supernatant samples were run on 4-12% NuPAGE gels (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions to check for expression. The gel was stained with Simply Blue stain (Invitrogen, Carlsbad, CA). Purification of GH43, GH51 and GH3 enzymes from *T. reesei* fermentations was performed by cation exchange chromatography as described in Example 1.

8.2 Genes from other species

[0272] Pf43A and Pf51A sequences were obtained by sequencing of the *Penicillium funiculosum* genome. The queries used for the searches of the *P. funiculosum* genome were other fungal GH43 and GH51 homologs available in the public genomes. A gene prediction program named Augustus (<http://augustus.gobics.de/>) was used to verify the intron sequences, along with the start codon. Fv43D sequence was used to query the *Gibberella zeae* (*Fusarium graminearum*) and *Fusarium oxysporum* genomes available in the Broad Institute database and retrieve sequences for Gz43A and Fo43A respectively. The genes for Gz43A and Fo43A were synthesized by GeneArt (Geneart GmbH, Regensburg, Germany) with the CBH1 signal sequence in place of the native signal sequence. Neither gene contained introns. The Pf51A gene was codon-optimized and synthesized by GeneArt with the CBH1 signal sequence in place of the native signal sequence.

[0273] Pf43A and Pf51A were cloned and expressed in *Trichoderma reesei* using the following primers:

Pf43A:MH151 (5'-CACCATGCTTCAGCGATTTGCTTATATTTTACC-3') (SEQ ID NO:78)

Pf43A:MH152 (5'-TTATGCGAACTGCCAATAATCAAAGTTG-3') (SEQ ID NO:79)

Pf51A:SK1168 (5'-CACCATGTACCGGAAGCTCGCCGTG-3') (SEQ ID NO:80)

Pf51A:SK1169: (5'-CTACTCCGTCTTCAGCACAGCCAC-3') (SEQ ID NO:81)

[0274] Three genes were cloned from *Aspergillus fumigatus* for expression in *Trichoderma reesei*: GH11 xylanase 2 (AfuXyn2), GH11 xylanase 5 (AfuXyn5), and GH43 Af43A. The primers for AfuXyn2, AfuXyn5, and Af43A cloning primers are shown below:

AfuXyn2:

A.fumi-Q4WG11-F: (5'-CCGCGGCCGCGACCATGGTTTCTTTCTCCTACCTGCTGCTG-3') (SEQ ID NO:82)

A.fumi-Q4WG11-R:

(5'-CCGGCGCGCCCTTACTAGTAGACAGTGATGGAAGCAGATCCG-3') (SEQ ID NO:83)

AfuXyn5:

A.fumi-Q4WFZ8-F: (5'-CCGCGGCCGCGACCATGATCTCCATTTCTCGCTCAGCT-3') (SEQ ID NO:84)

A.fumi-Q4WFZ8-R:

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

Af43A:

SK1203: (5'-CACCATGGCAGCTCCAAGTTTATCC-3') (SEQ ID NO:86)

SK1204 - (5' TCAGTAGCTCGGGACCACTC-3') (SEQ ID NO:87)

[0275] The methods used for cloning and expression of all these genes were similar to the procedure described for cloning of the *Fusarium* genes. Additional genes including those listed in Table 1B, were cloned in a similar manner.

9. EXAMPLE 3: TESTING FOR ACTIVITY OF NOVEL HEMICELLULASES ON SYNTHETIC SUBSTRATES

[0276] The activities of Fv3A, Fv43A, Fv43B, Fv43D and a number of other proteins in, for

example, Table 2, were tested on synthetic substrates pNPX and pNPA as described in Synthetic substrate activity assay in Example 1. *T. reesei* Bxl1 was used in at 0.7 g/L. The other enzyme samples, and the Quad-delete host control, were added by volume from growth cultures (microtiter plate or shake flask scale). Therefore, the absolute activity cannot be compared across samples but is an indication of an active expressed protein with the relative pNPX and pNPA activity shown in Table 2. Activity on para-nitrophenyl substrates is not used as a predictor of performance in biomass saccharification.

10. EXAMPLE 4: HYDROLYSIS OF PRETREATED CORNCOB BY CELLULASE AND HEMICELLULASE PREPARATIONS

10.1 Saccharification performance of expressed proteins

[0277] The saccharification performance of expressed proteins as additions to an enzyme mixture with an L- α -arabinofuranosidase deficiency was evaluated. L- α -arabinofuranosidase candidates were evaluated in a 4-day cob saccharification assay by addition to an enzyme mixture of Accellerase® 1500/*T. reesei* Xyn3/Fv3A. The screen was conducted as described in the corncob Saccharification Assay (Example 1) with the following enzymes and amounts/concentrations:

- Accellerase® 1500, TP (Total Nitrogen) 54.2 mg/mL
- *Trichoderma reesei* Xyn3, 2.9 mg/mL TP (purified)
- Fv3A, 3.2 mg/mL TP (purified)
- Fv51A, 7.8 mg/mL TP (purified)
- Mg51A, 6.8 mg/mL TP (TCA/BCA)
- At51A, 6.7 mg/mL TP (TCA/BCA)
- Pt51A, 3.3 mg/mL TP (TCA/BCA)
- Ss51A, 3.0 mg/mL TP (TCA/BCA)
- Vd51A, 6.8 mg/mL TP (TCA/BCA)
- Cg51B, 3.6 mg/mL TP (TCA/BCA)
- Af43A, 2.6 mg/mL TP (TCA/BCA)
- Pf43A, 2 mg/mL TP (TCA/BCA)
- Fv43E, 1.37 mg/mL TP (TCA/BCA)

[0278] The total protein (TP) of purified samples was determined by A280 unless otherwise noted. The total protein of unpurified samples was determined by BCA according to manufacturer instructions, unless otherwise noted. Accellerase® 1500 was added at 20 mg protein/g cellulose; *Trichoderma reesei* Xyn3 was added at 5 mg protein/g cellulose; Fv3A was added at 5 mg protein/g cellulose. Fv51A, Mg51A, At51A, Pt51A, Ss51A, Vd51A, Cg51 B, Af43A, Pf43A, or Fv43E were added at 1, 3, and 5 mg protein/g cellulose. Following 4 days

incubation, 50°C, 200 rpm, the assay plate was quenched and analyzed by HPLC for soluble sugars.

[0279] Enzymes were found that enhanced glucose or xylose or arabinose yield, or reduced cellobiose or xylobiose concentration in this enzyme mixture. Results are shown in Figures 30A and 30B.

[0280] The saccharification performance of expressed proteins was also evaluated as additions to the enzyme mixture with an L- α -arabinofuranosidase and β -xylosidase deficiency. β -xylosidase candidates were evaluated in a 3 day cob saccharification assay by addition to an enzyme mixture of Accellerase® 1500/*T. reesei* Xyn3. The screen was conducted as described in the corncob saccharification assay (Example 1) with the following enzymes and amounts/concentrations:

- Accellerase® 1500, TP (total nitrogen) 54.2 mg/mL
- *Trichoderma reesei* Xyn3, 2.9 mg/mL (purified)
- Fv3A, 3.2 mg/mL (purified)
- Fv43D, 6.8 mg/mL (purified)
- Pf43A, 2 mg/mL TP (BCA)
- Pf43B, 2.7 mg/mL TP (BCA)
- Fv43E, 1.37 mg/mL TP (BCA)
- Fv43F, 2.8 mg/mL TP (BCA)
- Fv30A, 2.7 mg/mL TP (BCA)

[0281] Accellerase® 1500 was added at 17.9 mg protein/g cellulose; *T. reesei* Xyn3 was added at 5 mg protein/g cellulose. Fv3A, Fv43D, Pf43A, Pf43B, Fv43E, Fv43F, or Fv30A were added at 1, 3, and 5 mg protein/g cellulose (Fv43E was only added at 1 and 3 mg/g). Following 3 days incubation, 50°C, 200 rpm, the assay plate was quenched and analyzed by HPLC for soluble sugars. Results are shown in Figure 31. Enzymes were found that enhanced glucose, or xylose or arabinose yield, or reduced cellobiose or xylobiose concentration in this enzyme mixture.

[0282] The saccharification performance of expressed proteins was also evaluated as additions to the enzyme mix with a xylanase deficiency. During the construction of the *Trichoderma reesei* integrated expression strains (described in Example 9 below), one *T. reesei* strain (strain # 44) was isolated that over-expressed Bgl1, Fv3A, Fv51A, Fv43D proteins but did not over-express endo-xylanase. This strain was used as the background to which candidate xylanases were added for performance screening. Endo-xylanase candidates were evaluated in a 3 day cob saccharification assay by addition to the enzyme products from strain #44. The screen was conducted as described in the corncob saccharification assay (Example 1) with the following enzymes and amounts/concentrations:

- Strain #44 enzyme product 78.6 mg/mL TP (modified Biuret)

- *Trichoderma reesei* Xyn3 2.9 mg/mL TP (purified)
- AfuXyn2 3.3 mg/mL TP (purified)
- AfuXyn3 5.8 mg/mL TP (purified)
- AfuXyn5 14.8 mg/mL TP (purified)
- PfuXyn1 1.9 mg/mL TP (purified)
- SspXyn1 1.2 mg/mL TP (purified)

[0283] The enzyme composition produced by Strain #44 was added at 20 mg protein/g cellulose; candidate xylanase enzymes were added at 3 and 7 mg protein/g cellulose. Following 3 days incubation, at 50°C, 200 rpm, the assay plate was quenched and analyzed by HPLC for soluble sugars. Enzymes were found that enhanced xylose, or glucose, or arabinose yield, or reduced cellobiose or xylobiose concentration in this enzyme mixture. Results are shown in Figure 32.

[0284] The saccharification performance of expressed Fv51A and Pa51A proteins in enzyme mixtures with L- α -arabinofuranosidase deficiency was evaluated and compared. Fv51A and Pa51A were evaluated in a 3 day cob saccharification assay by addition to an enzyme mixture of Accellerase® 1000/*T. reesei* Xyn2/Bxl1 or Fv3A. The screen was conducted as described in the corncob saccharification assay (Example 1) with the following enzymes and amounts/concentrations:

- Accellerase® 1000, 60.6 mg/mL TP (total nitrogen)
- *Trichoderma reesei* Xyn2 4.1 mg/mL TP (purified)
- *Trichoderma reesei* Bxl1 69 mg/mL TP (TCA/total nitrogen)
- Fv3A 65 mg/mL TP (TCA/total nitrogen)
- Fv51A 43 mg/mL TP (TCA/BCA)
- Pa51A 85.4 mg/mL TP (TCA/total nitrogen)

[0285] Accellerase® 1000 was added at 20 mg protein/g cellulose; *Trichoderma reesei* Xyn2 was added at 5 mg protein/g cellulose; *Trichoderma reesei* Bxl1 or Fv3A was added at 5 mg protein/g cellulose. Fv51A was added at 5 mg protein/g cellulose. Pa51A was added at 1, 2, or 5 mg protein/g cellulose. Following 3 days incubation, at 50°C, 200 rpm, the assay plate was quenched and analyzed by HPLC for soluble sugars. Enzyme combinations were found that enhanced xylose, or glucose, or arabinose yield, or reduced cellobiose or xylobiose concentration in this enzyme mixture. Results are shown in Figure 33.

[0286] Fv51A and Pa51A also were evaluated in a 3 day cob saccharification assay by addition to an enzyme mixture of Accellerase® 1000/ *Trichoderma reesei* Xyn2. Purified Fv51A (29 mg/mL TP) and Pa51A (29 mg/mL) were used in this part of the study. Accellerase® 1000 was added at 17.5 mg protein/g cellulose; *Trichoderma reesei* Xyn2 was added at 4.4 mg protein/g cellulose. Fv51A was added at 4.4 mg protein/g cellulose. Pa51A was added at 0.9, 1.8, and

4.4 mg protein/g cellulose. Following 3 days incubation, at 50°C, 200 rpm, the assay plate was quenched and analyzed by HPLC for soluble sugars. Enzyme combinations were found that enhanced xylose, or glucose, or arabinose yield in this enzyme mixture. Results are shown in Figure 34.

[0287] The saccharification performance of expressed proteins was also evaluated as additions to the enzyme mix with β -xylosidase deficiency. β -xylosidase candidates were evaluated in a 3 day cob saccharification assay by addition to an enzyme mixture of Accellerase® 1000/ *Trichoderma reesei* Xyn2/Fv51A. The screen was conducted as described in the corncob saccharification assay (Example 1) with the following enzymes and amounts/concentrations:

- Accellerase® 1000, TP (TCA/Total nitrogen) 60.6 mg/mL
- *Trichoderma reesei* Xyn2, 4.1 mg/mL TP (purified)
- Fv3A, 65 mg/mL TP (TCA/Total nitrogen)
- Fv3B, 62.9 mg/mLTP (TCA/Total nitrogen)
- Fv39A, 47.5 mg/mL TP (TCA/Total nitrogen)
- Fv30B, 62.9 mg/mL TP (TCA/Total nitrogen)
- Fv51A, 43 mg/mL TP (TCA/BCA)

[0288] Accellerase® 1000 was added at 20 mg protein/g cellulose; *Trichoderma reesei* Xyn2 was added at 5 mg protein/g cellulose; Fv51A was added at 5 mg protein/g cellulose. Fv39A, Fv30B, Fv3A, or Fv3B were added at 1, 2, or 5 mg protein/g cellulose. Following 3 days incubation, at 50°C, 200 rpm, the assay plate was quenched and analyzed by HPLC for soluble sugars. Enzymes and combinations were found that enhanced glucose or xylose or arabinose yield, or reduced cellobiose or xylobiose concentration in this enzyme mixture, with or without another β -xylosidase (*T. reesei* Bxl1). Results are shown in Figures 35A-35C.

10.2 Activity of candidate endo-xylanases with birchwood xylan

[0289] The activity of candidate endo-xylanases was evaluated using birchwood xylan as a substrate using the following assay. Ninety microliters of a 1% (wt/vol) birchwood xylan (Sigma X0502) stock solution was added to wells in a 96-well microtiter plate and pre-incubated at 50°C for 10 min. Enzyme dilutions and xylose standards were added (10 μ L) to the microtiter plate and the plates incubated at 50°C for 10 min. Meanwhile, 100 μ L of DNS solution was added to PCR tubes. Following the 10-min incubation, 60 μ L of the enzyme reaction was transferred to the PCR tubes containing the DNS solution. The tubes were incubated in a thermocycler at 95°C for 5 min, and then cooled to 4°C. One hundred microliters of the reaction mixture was transferred to a 96-well plate, and absorbance at 540 nm was measured. A xylose standard curve was generated and used to calculate the activity. One xylanase unit is defined as the amount of enzyme required to generate 1 μ mole of xylose reducing sugar

equivalents per minute under the conditions of the assay. Results are shown in Table 3.

10.3 Enzyme hydrolysis of arabinoxylan oligomers from saccharified corncob

[0290] In this study, enzyme hydrolysis of the arabinoxylan oligomers remaining after digestion of dilute ammonia pretreated corncob with cellulase and hemicellulase preparations was monitored. Preparation of crude oligomers is described in Example 1. Total oligomer sugars were determined by HPLC (see Example 1) after acid hydrolysis of the crude oligomers with 2% (v/v) sulfuric acid in a sealed vial at 121 °C for 30 min. The sugar concentrations were corrected for a small amount of sugar degradation as determined by control samples of known sugar mixtures treated by the same procedure. The concentration of total sugars in the crude oligomer preparation determined by this method was 45 g/L glucose, 168 g/L xylose, and 46 g/L arabinose. When accounting for monomer sugar present in crude oligomers before acid hydrolysis, 86% of the glucose, 90% of the xylose, and 43% of the arabinose was present in oligomeric form. Various β -xylosidases, arabinofuranosidases, and mixtures thereof were tested for increased conversion of arabinose monomer from the crude oligomers preparation. The crude oligomer preparation was diluted 20-fold to 12 g/L oligomers in 50 mM Sodium Acetate buffer, pH 5.0, and maintained at 50°C in a heating block in capped 1.5 mL Eppendorf tubes. Enzymes were added at final concentrations of 0.06-0.09 g/L and incubated for 24 h to reach completion. Samples were then removed for HPLC analysis of monomer sugars as described in Example 1. The results are listed in Table 4 as % conversion to monomer sugar based on total sugar as determined by acid hydrolysis.

[0291] To obtain the highest yields (44-71%) of arabinose from the remaining arabinoxylan oligomers in the crude oligomer mix, the data in Table 4 show that binary combinations of Fv3A+Fv51A, Fv3A+Fv43B, and Fv43A+Fv43B provide the best results. From the sequence families and activity on artificial substrates, it is deduced that Fv3A is a β -xylosidase and Fv51A is an L- α -arabinofuranosidase.

[0292] It is known that arabinose sugars in arabinoxylan from corncob are frequently linked to xylose at both the 2 and 3 carbon positions of the arabinose sugar. Thus the activity of Fv3A is likely to hydrolyze the xyl(1-2)ara linkage which then makes available the ara(1-3) xyl linkage to hydrolysis by the L- α -arabinofuranosidase. Of the β -xylosidases tested only Fv3A and Fv43A appeared to have this activity. Also in this screening only Fv43B appeared to have L- α -arabinofuranosidase activity among the Family 43 members from *Fusarium verticillioides*. Results are shown in Table 4.

11. EXAMPLE 5: SUBSTRATE RANGE OF β -XYLOSIDASES FOR EFFECTIVE CORNCOB HYDROLYSIS

[0293] In this example, the substrate range of 3 β -xylosidases and their relation to effective

conversion of corncob xylooligomers to monomer sugars were determined. Preparation of corncob hydrolysate containing oligomeric sugars and assay of monomer sugars was performed as described in Example 1. The proton NMR spectra of oligomeric sugars with degree of polymerization (DP) greater than 2 as separated by size exclusion chromatography on Bio-Gel P2 were determined (Figure 36 and Figure 37). The spectra of oligomers before enzyme treatment is labeled "MD07 oligomers" in the bottom panel of Figure 36 and spectra of the same oligo containing fractions after enzyme treatments are in the remaining panels of Figure 36 and Figure 37 labeled with the treatment enzyme.

[0294] The Bio-gel P2 fractions containing oligomers of greater than DP2 (5-10 mg) were lyophilized then dissolved in 0.7 mL of a D₂O solution containing 0.5 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as internal standard. Aliquots of 0.55 mL were used for the NMR samples in order to optimize suppression of the residual water peak. Standard versions of the Varian 2D correlation pulse sequences were used, with optimization of the ¹³C spectral width for the heteronuclear experiments. Spectra were acquired on a Varian Unity Inova using a high sensitivity cryoprobe operating at 500 MHz. Structure elucidation was done by identifying the correlations that characterize the spin-system for the individual sugar residues and then identifying the inter-glycosidic correlations.

[0295] The arabinose containing oligomers as determined by NMR, are dominated by one or more branched structures in which arabinose is linked β -1 $\rightarrow$ 3 to a xylose residue in a polymer fragment. The arabinose residue in the resulting branch is further substituted by a xylose residue linked β -1 $\rightarrow$ 2 to the arabinose. Little arabinose without this second substitution is present in the remaining xylo-oligomers. *T. reesei* Bxl1 is not very effective at cleaving the furthest out xylose 1 $\rightarrow$ 2 to arabinose bond as evidenced by the remaining signals at 5.5 to 5.55 ppm in the spectra. The combination of the Fv43A that is effective on longer chain xylose oligomers and the Fv43B L- α -arabinofuranosidase removes most but not all of the branched species with signals in the 5.5 ppm range and unlike treatment with the *T. reesei* Bxl1, leaves none of the signal at 5.35 that is attributable to remaining arabinose branched to xylose oligomer.

[0296] The anomeric proton region of the spectra of the remaining xylo-oligomers after treatment with the Fv51A alone or in combination with the Fv3A show different results. The L- α -arabinofuranosidase alone is not effective at reducing the complexity of signals in the region. The Fv3A β -xylosidase removes essentially all of the xylose subtending the arabinose branch leaving mostly simply linked arabinose 1 $\rightarrow$ 3 xylose oligomer structures. Addition of the L- α -arabinofuranosidase to the β -xylosidase results in a fairly complete conversion to monomer sugars as evidenced by the increase in signal coming from the α and β anomeric protons of reducing sugars.

[0297] It can thus be concluded that the increased effectiveness of one β -xylosidase over another is likely to be due to the substrate range in terms of the structural complexity of the aglycon allowable as substrate.

12. EXAMPLE 6: SEQUENCE COMPARISON AND CRITICAL RESIDUES IN DETERMINING SUBSTRATE RANGE IN GH3 B-XYLOSIDASES

[0298] The *T. reesei* β -xylosidase (Bxl1) is 61% similar to the *F. verticillioides* ortholog (Fv3A) and shares 42% sequence identity as shown in the alignment of Figure 38. Fv3A has much broader substrate range than Bxl1, which is likely to be attributable to the non-conserved amino acids among the two proteins.

13. EXAMPLE 7: DETERMINATION OF DEFINED HEMICELLULASE ACTIVITY FOR CORNCOB HYDROLYSIS TO MONOMER SUGARS

[0299] Pretreated corncob as described in Example 1 was used as water slurry adjusted to about pH 5 with H₂SO₄ in water at 18.6 g dry corncob solids/100 g of total slurry. 0.78 g of the slurry was added to 4 mL glass vials and sufficient pH 5.0, 50 mM Sodium Acetate buffer was added to give a total reaction weight of 1.06 g after the desired enzyme additions. The hemicellulases were added as the purified preparations described in Example 1. Supernatant from the quad deleted *T. reesei* strain (Quad in Tables 5 and 6) is the concentrate of background proteins expressed by the *T. reesei* strain deleted in 4 major cellulase activities (described in WO 05/001036). Accellerase® 1000 is a whole cellulase mixture with high β -glucosidase activity. The vials were incubated at 230 rpm in an orbital shaker at 48°C for 72 h then 2 mL of water was added. A sub-sample was taken and further diluted, centrifuged and filtered for HPLC analysis for monomer sugars as described in Example 1. Experimental results defining useful amounts of defined hemicellulase activity for hydrolyzing pretreated corncob to monomer sugars is shown in Table 5.

[0300] The results from the design-of-experiments (DoE) were fit to a surface model and used to determine best ratios of the 7 enzyme components for best yield of glucose, xylose and arabinose at the two total protein concentrations tested. Results of the ratios for the seven enzyme components are shown in Table 6.

[0301] Another exploration of the ratios (Table 7) was conducted including Fv3A, and again including Fv43D and holding that activity constant at a low level. The reaction set up and reaction conditions were identical to those described for the full DoE experiment.

[0302] Reactions (run numbers) 20, 21, and 22 contain only the whole cellulase enzyme mix and at the 21 mg/g of glucan loading monomerized about 48% of the glucose present in the cob and 24% of the xylose. Addition of the endoxylanase, *Trichoderma reesei* Xyn3 allowed decrease of the whole cellulase protein load while retaining about the same glucose monomer yield and increased the xylose monomer yield to about 40% (run# 1). All the combinations that gave arabinose yields of above 40% required the combination of Fv43D, Fv43A and Fv43B or Fv51A or of Fv3A and Fv51A or Fv43B. Those combinations also tended to have highest

release of xylose to monomer sugar.

[0303] Another set of reactions aimed at refining the required mix of hemicellulases was run holding both the loading of whole cellulase constant and the loading of the endoxylanase constant. Accellerase® 1000 whole cellulase preparation was held constant at 12 mg/g glucan and purified *T. reesei* Xyn3 endoxylanase was held at 6 mg/g xylan. Fv51A was the only L- α -arabinofuranosidase in the mixture, at different doses. Other reaction conditions remained the same but the hydrolysate was analyzed by size exclusion chromatography as described in Example 1. The quantitation of individual sugars was performed by peak area only and the results are shown in Table 8.

[0304] All combinations released similar amounts of glucose and about the same amount of total soluble xylose. The degree of reduction to monomer xylose varied by treatment. Without added activity to convert oligomer to monomer about 50% of the solubilized xylose remained oligomeric unless at least a β -xylosidase was added. Fv3A at 2 mg Fv3A protein/g xylan reduced >DP2 oligomers to 3.6 mg/mL. Addition of 2 mg Fv51A protein/g xylan to the 2 mg Fv3A protein/g xylan further decreased the >DP2 oligomers to 1.5 mg/mL. About 2 mg of Fv51A/g xylan appeared to be sufficient to reduce the >DP2 oligomers to a minimum when the required 2 mg/g Fv3A was present (Table 8).

[0305] A mix of 6 mg/g xylan *T. reesei* Xyn3, 2 mg/g Fv3A and either 1 or 2 mg/g Fv51A is a suitable loading to reduce total cob arabinoxylan to monomer sugars. The addition of Fv43D to the mix aids in taking the xylobiose or other DP2 oligomers to monomer. Arabinose hydrolysis to monomer was not measured in this experiment.

14. EXAMPLE 8: EFFECTIVENESS OF HEMICELLULASES AT PRODUCING MONOMER SUGARS FROM CORNCOB

[0306] In this example, the effectiveness of a set of purified hemicellulase activities at producing monomer xylose and arabinose sugars when acting alone on diluted ammonia pretreated corncob is demonstrated. Three mixtures (Mixes A, B, & C) of purified hemicellulases were prepared and used to hydrolyze hemicellulose in pretreated cob in 1 g total, 14% solids reactions prepared as in Example 1 and run under the conditions described in Example 1. Monomer sugars were analyzed by HPLC as described in Example 1 after 72 h of reaction and the amounts obtained are shown in Table 9.

- **Mix A:** 6 mg *Trichoderma reesei* Xyn3; 4 mg Fv3A; 1 mg Fv51A/g xylan
- **Mix B:** 6 mg *Trichoderma reesei* Xyn3; 1 mg Fv43D; 3 mg Fv43A; 3 mg Fv43B/g xylan
- **Mix C:** 6 mg *Trichoderma reesei* Xyn3; 3 mg Fv3A; 1 mg Fv43D; 1 mg Fv51A/g xylan

[0307] In this experiment the defined hemicellulase sets yielded slightly less monomer sugar

than seen in earlier experiments which included activities to solubilize cellulose. The yields were still greater than those seen with endoxylanase-only addition to whole cellulase preparations. The hemicellulase activities are effective in taking xylan to monomer.

[0308] The same set of mixtures was used on hemicellulose preparations made from corncob, total stover from grain sorghum, switchgrass and sugar cane bagasse using the procedures in the general methods in accordance with Example 1. Stock suspensions of each hemicellulose preparation at 100 mg/mL in 50 mM pH 5.0 Sodium Acetate buffer were made and the pH was checked. Each of them was diluted to 10 mg preparation per mL with more 50 mM acetate buffer. Aliquots of each enzyme mixture were added to 100 μ L of the 10 mg/mL suspension and the reactions were run in duplicate, incubated at 48°C for 6 h with agitation. Reactions were diluted with 100 μ L of water, centrifuged and filtered before HPLC analysis for monomer sugars as described in Example 1. 200 μ L of each hemicellulose suspension was diluted with 200 μ L of 0.8 N H₂SO₄, autoclaved at 121°C for 30 min on liquid cycle then filtered and sugars analyzed by HPLC as described in Example 1. Results shown in Table 10 are reported as the average monomer sugar released by the enzyme mixture as a percentage of the acid hydrolysable sugar present in the reaction.

[0309] Mixes A and C performed well on hemicellulose from cob and as seen in other experiments on whole pretreated cob. Mixtures containing Fv3A increase conversion of arabinose to monomer and give a slight advantage in conversion of xylose to monomer. All 3 mixtures work well on hemicellulose purified from other monocots. The mixing of one endoxylanase, either one or two β -xylosidases, one of which has substrate specificity beyond two or three xylose units linked β 1 $\rightarrow$ 4 and an L α -arabinofuranosidase results in an effective hemicellulase blend against monocot hemicellulose.

15. EXAMPLE 9: CONSTRUCTION OF THE INTEGRATED EXPRESSION STRAIN OF *TRICHODERMA REESEI*

[0310] An integrated expression strain of *Trichoderma reesei* was constructed that co-expressed five genes: *T. reesei* β -glucosidase gene *bgl1*, *T. reesei* endoxylanase gene *xyn3*, *F. verticillioides* β -xylosidase gene *fv3A*, *F. verticillioides* β -xylosidase gene *fv43D*, and *F. verticillioides* α -arabinofuranosidase gene *fv51A*.

[0311] The construction of the expression cassettes for these different genes and the transformation of *T. reesei* strain are described below.

15.1 Construction of the β -glucosidase expression cassette

[0312] The N-terminal portion of the native *T. reesei* β -glucosidase gene *bgl1* was codon optimized by DNA 2.0 (Menlo Park, USA). This synthesized portion comprised of the first 447

bases of the coding region. This fragment was PCR amplified using primers SK943 and SK941. The remaining region of the native *bgl1* gene was PCR amplified from a genomic DNA sample extracted from *T. reesei* strain RL-P37, using primer SK940 and SK942. These two PCR fragments of the *bgl1* gene were fused together in a fusion PCR reaction, using primers SK943 and SK942:

Forward Primer SK943:

(5'- CACCATGAGATATAGAACAGCTGCCGCT-3') (SEQ ID NO:88)

Reverse Primer SK941:

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

Forward Primer SK940:

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

Reverse Primer SK942:

(5'- CCTACGCTACCGACAGAGTG-3') (SEQ ID NO:91)

[0313] The resulting fusion PCR fragments were cloned into the Gateway® Entry vector pENTR™/D-TOPO®, and transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen) resulting in the intermediate vector, pENTRY-943/942 (Figure 39). The nucleotide sequence of the inserted DNA was determined. The pENTRY-943/942 vector with the correct *bgl1* sequence was recombined with pTrex3g using a LR clonase® reaction protocol outlined by Invitrogen. The LR clonase reaction mixture was transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen), resulting in the final expression vector, pTrex3g 943/942 (Figure 40). The vector also contains the *Aspergillus nidulans amdS* gene encoding acetamidase as a selectable marker for transformation of *T. reesei*. The expression cassette was PCR amplified with primers SK745 and SK771 to generate product for transformation of the strain, using the electroporation method described in WO 08153712.

Forward Primer SK771: (5' - GTCTAGACTGGAAACGCAAC -3') (SEQ ID NO:94)

Reverse Primer SK745: (5' - GAGTTGTGAAGTCGGTAATCC -3') (SEQ ID NO:95)

15.2 Construction of the endoxylanase expression cassette

[0314] The native *T. reesei* endoxylanase gene *xyn3* was PCR amplified from a genomic DNA sample extracted from *T. reesei*, using primers xyn3F-2 and xyn3R-2.

Forward Primer xyn3F-2:

(5'-CACCATGAAAGCAAACGTCATCTTGTGCCTCCTGG-3') (SEQ ID NO:94)

Reverse Primer (xyn3R-2):

(5'-CTATTGTAAGATGCCAACAATGCTGTTATATGCCGGCTTGGGG-3') (SEQ ID NO:95)

[0315] The resulting PCR fragments were cloned into the Gateway® Entry vector pENTR™/D-TOPO®, and transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen) resulting in the intermediate vector, pENTR/Xyn3 (Figure 41). The nucleotide sequence of the inserted DNA was determined. The pENTR/Xyn3 vector with the correct *xyn3* sequence was recombined with pTrex3g using a LR clonase® reaction protocol outlined by Invitrogen. The LR clonase reaction mixture was transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen), resulting in the final expression vector, pTrex3g/Xyn3 (Figure 42). The vector also contains the *Aspergillus nidulans amdS* gene encoding acetamidase as a selectable marker for transformation of *T. reesei*. The expression cassette was PCR amplified with primers SK745 and SK822 to generate product for transformation of the strain, using the electroporation method.

Forward Primer SK745: (5' - GAGTTGTGAAGTCGGTAATCC-3') (SEQ ID NO:96)

Reverse Primer SK822: (5' - CACGAAGAGCGGCGATTTC-3') (SEQ ID NO:97)

15.3 Construction of the β -xylosidase *Fv3A* expression cassette

[0316] The *F. verticilloides* β -xylosidase *fv3A* gene was amplified from a *F. verticilloides* genomic DNA sample using the primers MH124 and MH125.

Forward Primer MH124: (5' - CAC CCA TGC TGC TCA ATC TTC AG -3') (SEQ ID NO:98)

Reverse Primer MH125: (5' - TTA CGC AGA CTT GGG GTC TTG AG -3') (SEQ ID NO:99)

[0317] The PCR fragments were cloned into the Gateway® Entry vector pENTR™/D-TOPO®, and transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen) resulting in the intermediate vector, pENTR-Fv3A (Figure 43). The nucleotide sequence of the inserted DNA was determined. The pENTRY-Fv3A vector with the correct *fv3A* sequence was recombined with pTrex6g using a LR clonase® reaction protocol outlined by Invitrogen. The LR clonase reaction mixture was transformed into *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen), resulting in the final expression vector, pTrex6g/Fv3A (Figure 44). The vector also contains a chlorimuron ethyl resistant mutant of the native *T. reesei* acetolactate synthase (*als*) gene, designated *alsR*, which is used together with its native promoter and terminator as a selectable marker for transformation of *T. reesei* (WO2008/039370 A1). The expression cassette was PCR amplified with primers SK1334,

SK1335 and SK1299 to generate product for transformation of *T. reesei*, using the electroporation method (see, e.g., WO2008153712 A2).

Forward Primer SK1334: (5' - GCTTGAGTGTATCGTGTAAG -3') (SEQ ID NO:100)

Forward Primer SK1335: (5' - GCAACGGCAAAGCCCCACTTC -3') (SEQ ID NO:101)

Reverse Primer SK1299: (5' - GTAGCGGCCGCCTCATCTCATCTCATCCATCC -3') (SEQ ID NO:102)

15.4 Construction of the β -xylosidase Fv43D expression cassette

[0318] For the construction of the *F. verticilloides* β -xylosidase Fv43D expression cassette, the *fv43D* gene product was amplified from *F. verticilloides* genomic DNA using the primers SK1322 and SK1297. A region of the promoter of the endoglucanase gene *egl1* was PCR amplified from *T. reesei* genomic DNA extracted from strain RL-P37, using the primers SK1236 and SK1321. These two PCR amplified DNA fragments were subsequently fused together in a fusion PCR reaction using the primers SK1236 and SK1297. The resulting fusion PCR fragment was cloned into pCR-Blunt II-TOPO vector (Invitrogen) to give the plasmid TOPO Blunt/Pegl1-Fv43D (Figure 45) and *E. coli One Shot® TOP10* Chemically Competent cells (Invitrogen) were transformed using this plasmid. Plasmid DNA was extracted from several *E. coli* clones and confirmed by restriction digest.

Forward Primer SK1322: (5'-CACCATGCAGCTCAAGTTTCTGTC-3') (SEQ ID NO:103)

Reverse Primer SK1297: (5'-GGTTACTAGTCAACTGCCCGTTCTGTAGCGAG-3') (SEQ ID NO:104)

Forward Primer SK1236: (5'-CATGCGATCGCGACGTTTTGGTCAGGTCTG-3') (SEQ ID NO:105)

Reverse Primer SK1321 : (5'-GACAGAACTTGAGCTGCATGGTGGGACAACAAGAAGG-3') (SEQ ID NO:106)

[0319] The expression cassette was PCR amplified from TOPO Blunt/Pegl1-Fv43D with primers SK1236 and SK1297 to generate product for transformation of *T. reesei*, using the electroporation method as described in WO2008153712A2.

15.5 Construction of the α -arabinofuranosidase expression cassette

[0320] For the construction of the *F. verticilloides* α -arabinofuranosidase gene *fv51A* expression cassette, the *fv51A* gene product was amplified from *F. verticilloides* genomic DNA using the primers SK1159 and SK1289. A region of the promoter of the endoglucanase gene *egl* was PCR amplified from *T. reesei* genomic DNA sample extracted from strain RL-P37, using the primers SK1236 and SK1262. These two PCR amplified DNA fragments were subsequently fused together in a fusion PCR reaction using the primers SK1236 and SK1289. The resulting fusion PCR fragment was cloned into pCR-Blunt II-TOPO vector (Invitrogen) to give the plasmid TOPO Blunt/Pegl1-Fv51A (Figure 46) and *E. coli* One Shot® TOP10 Chemically Competent cells (Invitrogen) were transformed using this plasmid.

Forward Primer SK1159: (5'-CACCATGGTTCGCTTCAGTTCAATCCTAG-3') (SEQ ID NO:107)

Reverse Primer SK1289: (5'-GTGGCTAGAAGATATCCAACAC-3') (SEQ ID NO:108)

Forward Primer SK1236: (5'-CATGCGATCGCGACGTTTTGGTCAGGTCG-3') (SEQ ID NO:109)

Reverse Primer SK1262:

(5'-GAACTGAAGCGAACCATGGTGTGGGACAACAAGAAGGAC-3') (SEQ ID NO:110)

[0321] The expression cassette was PCR amplified with primers SK1298 and SK1289 to generate product for transformation of *T. reesei* using the electroporation method.

Forward Primer SK1298: (5'-GTAGTTATGCGCATGCTAGAC-3') (SEQ ID NO:111)

Reverse Primer SK1289: (5'-GTGGCTAGAAGATATCCAACAC-3') (SEQ ID NO:112)

15.6 Co-Transformation of *T. reesei* with the β -glucosidase and endoxylanase expression cassettes

[0322] A *Trichoderma reesei* mutant strain, derived from RL-P37 (Sheir-Neiss, G et al. Appl. Microbiol. Biotechnol. 1984, 20:46-53) and selected for high cellulase production was co-transformed with the β -glucosidase expression cassette (*cbhl* promoter, *T. reesei* β -glucosidase1 gene, *cbhl* terminator, and *amdS* marker), and the endoxylanase expression cassette (*cbhl* promoter, *T. reesei* *xyn3*, and *cbhl* terminator) using PEG-mediated transformation (Penttila, M et al. Gene 1987, 61 (2):155-64). Numerous transformants were isolated and examined for β -glucosidase and endoxylanase production. One transformant called *T. reesei* strain #229 was used for transformation with the other expression cassettes.

15.7 Co-transformation of *T. reesei* strain #229 with two β -xylosidase and α -arabinofuranosidase expression cassettes

[0323] *T. reesei* strain #229 was co-transformed with the β -xylosidase *fv3A* expression cassette (*cbh1* promoter, *fv3A* gene, *cbh1* terminator, and *alsR* marker), the β -xylosidase *fv43D* expression cassette (*egl1* promoter, *fv43D* gene, native *fv43D* terminator), and the *fv51A* α -arabinofuranosidase expression cassette (*egl1* promoter, *fv51A* gene, *fv51A* native terminator) using electroporation. Transformants were selected on Vogels agar plates containing chlorimuron ethyl (80 ppm). Vogels agar was prepared as follows, per liter.

50 x Vogels Stock Solution (below)	20 mL
BBL Agar	20 g
With deionized H ₂ O bring to post-sterile addition:	980 mL
50% Glucose	20 mL

[0324] 50 x Vogels Stock Solution (WO 2005/001036), per liter:

In 750 mL deionized H₂O, dissolve successively:

Na ₃ Citrate*2H ₂ O	125.00 g
KH ₂ PO ₄ (Anhydrous)	250.00 g
NH ₄ NO ₃ (Anhydrous)	100 g
MgSO ₄ *7H ₂ O	10.00 g
CaCl ₂ *2H ₂ O	5.00 g
Vogels Trace Element Solution	5.0 mL
Vogels Biotin Solution	2.5 mL
With deionized H ₂ O,	bring to 1 L

[0325] Numerous transformants were isolated and examined for β -xylosidase and L- α -arabinofuranosidase production. Transformants were also screened for biomass conversion performance according to the cob saccharification assay described in Example 1. Examples of *T. reesei* integrated expression strains described herein are H3A, 39A, A10A, 11A, and G9A, which express all of the genes for *T. reesei* beta-glucosidase 1, *T. reesei* Xyn3, Fv3A, Fv51A, and Fv43D, at different ratios (Table 11). Examples of *T. reesei* integrated expression strains described herein also include 44A, 69A, G6A and 102, and each includes most of the genes for *T. reesei* beta-glucosidase 1, *T. reesei* XYN3, Fv3A, Fv51A, and Fv43D, expressed at different ratios. Strain 44A lacked overexpressed *T. reesei* XYN3; strain 69A lacked Fv51A (confirmed by Western Blot, not shown); strains G6A and 102 lacked Fv3A (Table 11), as determined by HPLC protein analysis (Example 1).

16. EXAMPLE 10: SACCHARIFICATION PERFORMANCE OF *T. REESEI* INTEGRATED EXPRESSION STRAINS ON AMMONIA PRETREATED CORNCOB

[0326] The saccharification performance of enzyme compositions produced by *T. reesei* integrated expression strains on dilute ammonia pretreated corncob was evaluated. *T. reesei* enzyme samples were generated as either ultrafiltration concentrates (UFC) or centrate. For the generation of UFC, *T. reesei* fermentation broths (14L-scale) were obtained after cell separation by centrifugation, concentrated using membrane-ultrafiltration through a Millipore 10 kD molecular weight cut off membrane. Then pH was adjusted to 4.8. The cell-separated broth was then polished by filtration, using FW6 Buchner filtration. Each enzyme sample was assayed for total protein concentration using the modified Biuret method.

[0327] The saccharification performance was evaluated in vials or in shake flasks. Each enzyme preparation was assayed for saccharification performance on 20% dry solids (DS) loading of dilute ammonia pretreated corncob (see, WO2006/110901). All saccharification reactions were then titrated with sulfuric acid to pH 5.0 and sodium azide was added to a final concentration of 0.01% (w/v), for microbial contamination control. Each saccharification reaction was then dosed with 20 mg of total protein (TP) enzyme preparation per g of substrate glucan or xylan, as appropriate. Accellerase® 1500 (Ac1500) and the integrated strain UFC's were dosed at 20 mg total protein/g glucan. An enzyme blend was prepared in a ratio of 25:9:4:3:1 Accellerase® 1500:Xyn3:Fv3A:Fv51A: FV43D. In the blend, Accellerase® 1500 was dosed at 20 mg total protein/g glucan and the hemicellulases were dosed per g xylan (4.3 mg Xyn3/g of xylan, 1.7 mg Fv3A/g of xylan, 1.4 mg Fv51A/g of xylan, 0.5 mg Fv43D/g of xylan). Each saccharification reaction was incubated at 50°C in a rotary shaker set to 200 rpm, then sampled and diluted 10 x (v/v) before monomeric sugar concentration was determined using HPLC analysis (detailed in Section 16.1 below under the "monomeric HPLC analysis" section) after 1, 2, 3 and 7 days of saccharification (Figures 47A, 47B). On day 3 of saccharification, each reaction was also sampled by weight (w/w) for oligomeric sugar concentration by HPLC analysis (Section 16.1, Figure 47C). These results show that the enzyme compositions produced by integrated strains H3A and G9A provide better glucose and xylose yields than Accellerase®1500 or enzyme blends created from individually expressed enzymes.

[0328] A chromatographic comparison of the enzyme composition produced by three different integrated strains is shown in Figure 47D.

16.1 HPLC Analysis

[0329] Monomeric Sugar HPLC Analysis: Each sample was analyzed by HPLC using a BioRad Aminex HPX-87H ion exclusion column (300 mm x 7.8 mm). All day 1, 2, 3, and 7 samples were diluted 10 x volumetrically with 5 mM sulfuric acid, filtered through a 0.2 µm filter before injection into the HPLC and run under manufacture specifications.

[0330] Oligomeric Sugar HPLC Analysis (acid hydrolysis): Day 3 saccharification samples were

diluted 10 x by weight with Milli-Q water, then sulfuric acid was added to the final concentration of 4% (w/w). A xylose and glucose standard (sugar recovery standard - "SRS") of known concentration was also prepared and measured for monomeric sugar HPLC analysis as stated above, and oligomer sugar HPLC analysis. Oligomer HPLC samples and were then autoclaved at 121 °C for 15 min, and filtered through a 0.2 µm filter before injection into the HPLC BioRad Aminex HPX-87H ion exclusion column (300 mm x 7.8 mm), and run under manufacturer specifications. Oligomer sugar concentration was determined by multiplying the percent retained xylose and glucose concentration of the sugar recovery standard (SRS) after acid hydrolysis, by the post acid hydrolysis HPLC sugar concentrations of each sample, then subtracting the monomeric sugar concentration determined in the "monomeric sugar HPLC Analysis" section above.

17. EXAMPLE 11: IDENTIFICATION OF PHYLOGENETICALLY BROAD ENZYME CLASSES WHICH CAN IMPROVE THE SACCHARIFICATION PERFORMANCE OF *T. REESEI* INTEGRATED EXPRESSION STRAIN

[0331] In this example, enzyme activities which limit the efficacy of enzyme compositions produced by a *T. reesei* integrated expression strain were identified and enzymes from different species which can compensate for those limiting activities or which could substitute for integrated strain components are exemplified.

[0332] 0.95g of pretreated corncob as described in Example 1 was added to 20 mL glass vials. Sufficient pH 5.0, 50 mM Sodium Acetate buffer and 1 N H₂SO₄ were added to give a total reaction weight of 3.00 g at 22 g dry corncob solids/100g of total slurry, pH 5.0 post enzyme additions. The *T. reesei* enzyme composition produced by integrated strain H3A (Figures 48A, 49A, and 50A) was added as an ultrafiltered (*i.e.*, cell-free) concentrate to 7 mg total protein per g of glucan and xylan combined in the feedstock to all vials. In addition, candidate hemicellulases were added as the ultrafiltered preparations from 14 L fermentation cultures expressed by *T. reesei* Quad delete strain (described in WO 05/001036). The candidate hemicellulases were added at 0, 0.5, 1.0 or 3.0 mg enzyme protein per g of glucan and xylan combined. The hemicellulases included constituents of the enzyme composition produced by the integrated strain itself including:

- *Fusarium verticillioides* Fv3A
- *Fusarium verticillioides* Fv51A
- *Fusarium verticillioides* Fv43D

[0333] As well as enzymes from different fungi

- *Fusarium oxysporum* Fo43A
- *Gibberella zeae* Gz43A
- *Penicillium funiculosum* Pf43A

- *Aspergillus fumigatus* Af43A
- *Podospora anserina* Pa51A
- *Penicillium funiculosum* Pf51A

[0334] The vials were incubated at 180 rpm in an orbital shaker at 48°C for 72 h. Then 12 mL of water was added. A sub-sample was taken and further diluted, centrifuged and filtered for HPLC analysis for monomer sugars as described in Example 1.

[0335] The results shown in Figure 48A and Figure 48B demonstrate that addition of β -xylosidase activities such as *Fusarium verticillioides* Fv43D and *Fusarium oxysporum* Fo43A markedly improved monomer xylose release relative to 7 mg *T. reesei* integrated strain total protein per g of Glucan and Xylan combined. Significant improvements were even observed at the low additions of 0.5 and 1.0 mg/g.

[0336] The results shown in Figure 49A and Figure 49B demonstrate that addition of all of the hemicellulases lead to a significant increase in monomer glucose, even >10% at hemicellulase loadings of 1-3 mg per g of Glucan and Xylan combined.

[0337] The results shown in Figure 50A and Figure 50B demonstrate that addition of GH51 enzymes such as *Fusarium verticillioides* Fv51A, and especially *Podospora anserine* Pa51A and *Penicillium funiculosum* Pfu51A, led to an increase in monomer arabinose level.

18. EXAMPLE 12: SACCARIFICATION OF VARIOUSLY PRETREATED SWITCHGRASS BY CELLULASE AND HEMICELLULASE PREPARATIONS

[0338] The saccharification performance of expressed cellulases and hemicellulases on pretreated raw switchgrass was evaluated. A range of conditions for dilute ammonia pretreatment of switchgrass were evaluated for saccharification performance with an enzyme cocktail composed of enzymes described herein. Pretreatment conditions vary and pretreatment efficacy affects enzymatic hydrolysis performance.

[0339] Pretreatment of raw switchgrass was performed in sealed, 6" x ½", stainless steel tubes that were immersed in a heated sand bath. A slurry of raw switchgrass, water, and ammonium hydroxide (-28% solution) was mixed to the desired percent solids and percent ammonia and then loaded into a pretreatment tube. Tubes were then held at the desired temperature (+/- 2°C) for the desired time and then quenched in ice water for approximately 1 min before being brought to room temperature. The pretreated slurry was removed from the tubes and allowed to dry overnight in the hood (>90% solids attainable).

[0340] Dried pretreated solids were then saccharified at 10% solids, pH 5, 50°C, 200 rpm using Accellerase® 1500, Xyn 3, Fv3A, Fv51A, and Fv43D (25, 9, 7, 3, 1 mg total protein/g

glucan or xylan respectively). A 5 mL total hydrolysate volume in 20 mL scintillation vials was used.

[0341] Glucan and xylan yields were based on monomeric glucose and xylose released compared to the glucan or xylan available from the raw biomass. Monomeric sugar concentrations were measured by HPLC (BioRad Aminex HPX 87-H column).

[0342] Pretreatment time, temperature, percent solids, and percent NH_3 , were varied over a wide range in order to optimize saccharification results. Each pretreatment condition that had both glucan and xylan yields better than -50% is considered a strong performer. The pretreatment parameters that performed strongly are listed in Table 12 along with their respective glucan, xylan, and total percent yields. Glucan and Xylan conversions are based on monomeric sugars released during saccharification as compared to glucan or xylan theoretically available in the raw switchgrass. Total conversion is glucose and xylose only (Figure 51).

19. EXAMPLE 13: SACCHARIFICATION OF PRETREATED SWITCHGRASS BY CELLULASE AND HEMICELLULASE PREPARATIONS

[0343] In addition to the above Examples, the saccharification performance of enzyme mixes and enzyme compositions produced by an integrated strain was tested on several substrates, pretreatments and conditions. These experiments show the range of performance using the enzyme mix or an integrated strain product. They demonstrate good performance across a range of substrates and pretreatments, pH, and temperatures.

[0344] Dilute ammonia pretreated switchgrass was prepared according to the methods and process ranges in WO06110901A : Switchgrass (38.7% glucan, 22.2% xylan, 2.5% arabinan, 23.2% lignin) was hammer-milled to pass through a 1mm screen, then pretreated at 160°C for 90 min with 6% NH_3 (weight/ weight DM, added as NH_4OH). This pretreated substrate was treated with enzyme mixes containing Accellerase® 1500, Multifect® Xylanase (both commercial products of Danisco A/S, Genencor Division, Palo Alto, CA), Fv3A, Fv51A, and Fv43D in a total reaction mass of 50g at 15% solids. The total protein (TP) of the commercial products was determined by Biuret assay. The other enzymes were ultra-filtration concentrates (UFCs) following expression in cellulase quad-deleted strains of *T. reesei*, with TP determined by Total Nitrogen analysis of TCA-precipitable protein. All reactions were dosed with Accellerase® 1500 at 25 mg TP/g Glucan and Multifect® Xylanase (MF Xyl) at 9 mg TP/g Xylan, and Fv3A, Fv51A, and Fv43D were added as indicated in Figure 56A-56B at 3.6 mg TP/g Xylan, 3.0 mg TP/g Xylan and 1.0 mg TP/g Xylan, respectively. All enzymes were dosed relative to the starting carbohydrate contents of the switchgrass before pretreatment. The saccharification reactions were carried out at 47°C and 33°C at pH 5.3 for three days.

[0345] The results at 33°C showed that addition of Fv3A, Fv51A, and Fv43D increased glucan

conversion (Figure 56A) and more than doubled the xylan conversion (Figure 56B). The results at 47°C showed that addition of Fv3A, Fv51A, and Fv43D gave some increased glucan conversion (Figure 56A) and more than doubled the xylan conversion (Figure 56B). Additions of Fv51A or Fv43D, alone, gave large increases in xylan conversion, especially to xylo-oligomers or xylose monomers, respectively. Addition of Fv3A alone increased xylose yields, but in combination with Fv51A gave a large increase in xylan conversion to monomer.

20. EXAMPLE 14: SACCHARIFICATION OF PRETREATED SWITCHGRASS BY AN INTEGRATED *T. REESEI* STRAIN

[0346] The saccharification performance of an enzyme composition produced by an integrated *T. reesei* strain (H3A) was evaluated on dilute ammonia pretreated switchgrass prepared according to the methods and process ranges in WO06110901A: Switchgrass (37% glucan, 21% xylan, 5% arabinan, 18% lignin) was hammer-milled to pass through a 1 mm screen, then pretreated at 160°C for 90 min with 10.0% NH₃ (weight / weight DM, added as NH₄OH). In duplicate 500 mL glass Erlenmeyer flasks, 50 g of pretreated slurry at 25% solids was saccharified at 48°C, pH 5.3 for 7 days, with supernatant from the integrated strain was dosed at 14 mg TP/g of carbohydrate (glucan plus xylan). TP of H3A was determined by Total Nitrogen analysis of TCA-precipitable protein. Enzymes were dosed relative to the starting carbohydrate contents of the switchgrass before pretreatment.

[0347] At the end of 7 days, high levels of glucan conversion (52-55%, Figure 57A) and xylan conversion (51%-53%, (Figure 57B) were measured by HPLC. These results show that the enzyme composition produced by the integrated strain can saccharify dilute ammonia pretreated switchgrass at high solids (25% dry matter).

21. EXAMPLE 15: SACCHARIFICATION OF HARDWOOD PULP BY AN INTEGRATED *T. REESEI* STRAIN

[0348] The saccharification performance of an enzyme composition produced by integrated *T. reesei* strain H3A was evaluated on industrial hardwood unbleached pulp (derived from Kraft process and oxygen delignification, Smurfit Kappa Cellulose Du Pin, Biganos, France) with the following composition: Glucan 75.1 %, Xylan 19.1 %, Acid soluble lignin 2.2 %. The enzymatic saccharification studies were carried out using NREL standard assay method LAP-009 "Enzymatic Saccharification of Lignocellulosic Biomass" (<http://www.nrel.gov/biomass/pdfs/42629.pdf>), except that the cellulose loading was different (varying from 9.3-20%) and a total mass of 100 g was used. The experimental condition was 200 rpm and pH 5.0, 50°C. Enzyme was dosed at 20 mg TP/g glucan (based on final dry matter) at the start of the experiment. Samples were taken at timed intervals if they were liquefied, and then analyzed by HPLC for sugar concentration. Glucose, xylose, and cellobiose concentration were determined using a Waters HPLC system (Alliance system, Waters Corp.,

Milford, MA). The HPLC column used for sugar analysis was from BioRad (Aminex HPX-87H ion exclusion column (300 mm x 7.8 mm), BioRad Inc., Hercules, CA). All samples were diluted 10 x with 5 mM sulfuric acid, filtered through a 0.2 µm filter before injection into the HPLC. As indicated (TABLE 13) two experiments were carried out in "fed-batch" mode: Pretreated hardwood pulp to an initial dry matter content of 7.0% at Time 0 and the rest of the substrate was added at discrete times in four equal portions during the first 24 h to bring the final dry solids loading to 20%.

[0349] Results showed high levels of glucan and xylan conversion to monomers (up to 89% and 90%, respectively) with the enzyme composition produced by integrated strain H3A. Conversions increased with higher enzyme loadings and longer saccharification times. Conversion was lower when the solids were at 20% than at 15%, but this lower conversion at 20% could be partially mitigated by using a fed-batch process.

22. EXAMPLE 16: SACCHARIFICATION OF HARDWOOD PULP BY AN INTEGRATED *T. REESEI* STRAIN OVER A RANGE OF TEMPERATURE AND PH

[0350] The saccharification performance of an enzyme composition produced by an integrated *T. reesei* strain (H3A) on industrial hardwood unbleached pulp (as used in the preceding example) at 7% cellulose loading and 20 mg TP/g glucan of integrated strain supernatant (based on final dry matter and composition of the pretreated substrate) was tested at temperatures from 45°C - 60°C and pH's from 4.65 - 5.4 (buffered in 0.1 M sodium citrate). Results after 2 days saccharification are shown in TABLE 14 and show good glucan and xylan conversions over the whole range of conditions tested. These results suggested optimum conditions for saccharification of this substrate as pH 4.9 and 50°C but good conversions are seen even at pH 5.0, 60°C. In a follow-up experiment at 50°C, including lower pH's (and otherwise unchanged experimental conditions) good saccharifications were seen at pH 3.8, pH 4.0 and pH 4.25, with glucose & xylose titers of 45.1 & 9.7 g/L; 50.0 & 11.4 g/L; and 57.2 & 13.2 g/L, respectively.

23. EXAMPLE 17: SACCHARIFICATION OF STEAM-EXPANDED SUGARCANE BAGASSE WITH AN ENZYME COMPOSITION PRODUCED BY AN INTEGRATED STRAIN AT DIFFERENT ENZYME DOSES

[0351] The saccharification performance of an enzyme composition produced by an integrated *T. reesei* strain (H3A) on steam-expanded sugarcane bagasse was evaluated at 7% cellulose loading. The bagasse was pretreated by steam injection in a StakeTech reactor at 210 psig, 200°C with a 4 min residence time. The pretreated material had the following composition: Glucan 40.9 %, Xylan 20.8 %, Lignin 27%. The integrated strain supernatant was dosed at 10, 20, 30, 50 or 80 mg TP/g glucan (based on final dry matter and composition of the pretreated substrate). Saccharification was carried out in a 5 mL reaction volume at 50°C, pH 5 for 3

days. The results (Figure 58A-C) showed that the integrated strain product out-performed Accellerase® 1500 (20 mg protein/g glucan) in glucan conversion to glucose (Figure 58A and Figure 58C) and, especially, xylan conversion to xylose (Figure 58B and Figure 58C) even at half the Accellerase® 1500 dose. Glucan conversion, especially at Day 1, increased significantly as the dose of H3A total protein increased (Figure 58A and Figure 58C), whereas xylan conversion was more rapid, with little increase from Day 1 to Day 3, and very high, even at the lowest dose of H3A total protein (Figure 58B and Figure 58C).

24. EXAMPLE 18: SACCHARIFICATION OF DILUTE-ACID PRETREATED CORN FIBER WITH VARIOUS ENZYMES

[0352] The saccharification performance of enzyme mixtures on dilute sulfuric acid pretreated corn fiber was evaluated in a 250 mL shake flask. In a typical experiment, corn fiber (initial composition 38% C6 sugars and 27% C5 sugars) was adjusted to 15% DS (dry solids) and 0.36% (w/w%) sulfuric acid was added. Corn fiber slurry was then autoclaved at 121 °C for 60 min. The slurry was then adjusted to pH 5.0 using 6 N NaOH. The sugar content of the pretreated sample was 21 g/L glucose and 12 g/L xylose. Enzymes were added to the pretreated substrate as follows (as indicated in Figure 59A, 59B, 59C); Accellerase® 1500 (AC 1500): 20 mg TP/g glucan and 45 mg TP/g glucan; Accellerase® 1500 + Multifect® Xylanase (MF): (25 mg/g glucan + 9mg TP/g xylan) and (25 mg TP/g glucan + 20 mg TP/g xylan); Accellerase® 1500 +Xyn 3+ Fv3A+ Fv51A+ Fv43D: (25 mg TP/g glucan + 9 mg TP/g xylan + 7 mg TP/g xylan + 3 mg TP/g xylan + 1 mg TP/g xylan) (the full enzyme "blend"). The total protein (TP) of the commercial products (Accellerase® 1500 and Multifect® Xylanase: Danisco US Inc., Genencor) were determined by Biuret assay. The other enzymes were ultra-filtration concentrates (UFCs) following expression in cellulase quad-deleted strains of *T. reesei* (as described earlier) - their TP were determined by Total Nitrogen analysis of TCA-precipitable protein. All enzymes were dosed relative to the starting carbohydrate contents of the corn fiber before pretreatment. Enzymatic saccharification was carried out at 200 rpm and 50°C for 24 h, 48 h and 120 h. Samples were withdrawn at different time intervals and analyzed for the formation of glucose and xylose sugars by HPLC. The adjusted sugar (glucose or xylose) reflected the sugar being produced from the enzymatic step, with the starting sugars subtracted.

[0353] The results show that the full enzyme blend out-performed the (Accellerase® 1500 + Multifect® Xylanase) ("AC 1500 + MF") in glucan conversion and in xylan conversion, even when both were dosed at the same total protein. The full enzyme blend gave almost complete glucan conversion of this substrate after 5 days saccharification.

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not

form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- [US2006050134A \[0007\]](#)
- [WO2008045977A \[0008\]](#)
- [US6660506B \[0134\]](#)
- [US6423145B \[0134\]](#)
- [US5536325A \[0135\]](#)
- [US6409841B \[0136\]](#)
- [US5705369A \[0137\]](#)
- [WO2004081185A \[0140\]](#)
- [US20070031918A \[0141\]](#)
- [WO06110901A \[0141\] \[0225\] \[0344\] \[0346\]](#)
- [WO2002095014A \[0165\]](#)
- [WO2004078919A \[0165\]](#)
- [US6022725A \[0165\]](#)
- [US6982159B \[0165\]](#)
- [US7045332B \[0165\]](#)
- [US7005289B \[0165\]](#)
- [US20060258554A \[0165\] \[0165\]](#)
- [US20070213249A \[0172\]](#)
- [US20090170181A \[0172\]](#)
- [US5536655A \[0172\]](#)
- [US20070111278A \[0172\]](#)
- [US20070148732A \[0172\]](#)
- [WO9117243A \[0172\]](#)
- [WO9117244A \[0172\]](#)
- [WO9110732A \[0172\]](#)
- [US6001639A \[0172\]](#)
- [US20070148730A \[0175\]](#)
- [US5405769A \[0187\]](#)
- [US6767728B \[0200\]](#)
- [US6426200B \[0200\]](#)
- [US6512110B \[0200\]](#)
- [US6254722B \[0200\]](#)
- [US6261828B \[0201\]](#)
- [US6077316A \[0201\]](#)
- [US6024766A \[0201\]](#)
- [US6021536A \[0201\]](#)

- [US6017751A \[0201\]](#)
- [US5980581A \[0201\]](#)
- [US20020142438A1 \[0201\]](#)
- [US6245546B \[0209\]](#)
- [US5709796A \[0213\]](#)
- [US6413928B \[0214\]](#)
- [US6399561B \[0214\]](#)
- [US6365561B \[0214\]](#)
- [US6380147B \[0214\]](#)
- [US7354755B \[0227\]](#)
- [WO05001036A \[0237\] \[0270\] \[0271\] \[0299\] \[0332\]](#)
- [WO2005001036A \[0259\] \[0324\]](#)
- [US20060003408A \[0271\]](#)
- [WO2009114380A \[0271\]](#)
- [US7713725B \[0271\]](#)
- [WO08153712A \[0313\]](#)
- [WO2008039370A1 \[0317\]](#)
- [WO2008153712A2 \[0317\] \[0319\]](#)
- [WO2006110901A \[0327\]](#)

Non-patent literature cited in the description

- BUNGAY, H. R. Energy: the biomass options Wiley 1981 0000 [\[0002\]](#)
- OLSSON LHAHN-HAGERDAL B Enzyme Microb Technol, 1996, vol. 18, 312-31 [\[0002\]](#)
- ZALDIVAR, J et al. Appl Microbiol Biotechnol, 2001, vol. 56, 17-34 [\[0002\]](#)
- GALBE, M et al. Appl Microbiol Biotechnol, 2002, vol. 59, 618-28 [\[0002\]](#)
- SORENSEN H R et al. BIOTECHNOLOGY AND BIOENGINEERING, 2003, vol. 81, 6726-731 [\[0009\]](#)
- SAHA B CAPPLIED MICROBIOLOGY AND BIOTECHNOLOGY, 2001, vol. 56, 762-766 [\[0010\]](#)
- TÖRRÖNEN et al. Biotechnology, 1992, vol. 10, 1461-65 [\[0041\]](#)
- MARGOLLES-CLARK et al. Appl. Environ. Microbiol., 1996, vol. 62, 103840-46 [\[0041\]](#)
- BARNETT et al. Bio-Technology, 1991, vol. 9, 6562-567 [\[0041\]](#)
- HENRISSAT et al. FEBS Lett, 1998, vol. 425, 2352-4 [\[0043\]](#)
- COUTINHO HENRISSAT Genetics, biochemistry and ecology of cellulose degradation T. Kimura. Tokyo, Uni Publishers Co 1999 0000 15-23 [\[0043\]](#)
- CANTAREL et al. Nucleic Acids Res., 2009, vol. 37, D233-38 [\[0044\]](#)
- LI et al. Biochem. J., 2001, vol. 355, 835-840 [\[0047\]](#)
- YANG et al. J. Mol. Biol., 2004, vol. 335, 1155-65 [\[0048\]](#)
- CZJZEK et al. J. Mol. Biol., 2005, vol. 353, 4838-46 [\[0048\]](#)

- NURIZZO et al. Nat. Struct. Biol., 2002, vol. 9, 9665-8 [0049]
- PONS et al. Proteins: Structure, Function and Bioinformatics, 2004, vol. 54, 424-432 [0049]
- BRUX et al. J. Mol. Bio., 2006, vol. 359, 97-109 [0049]
- HÖVEL et al. EMBO J., 2003, vol. 22, 194922-4932 [0050]
- PELL et al. J. Biol. Chem., 2004, vol. 279, 109597-9605 [0051]
- SCHMIDT et al. Biochem., 1999, vol. 38, 2403-2412 [0051]
- LO LEGGIO et al. FEBS Lett., 2001, vol. 509, 303-308 [0051]
- HAVUKAINEN et al. Biochem., 1996, vol. 35, 9617-24 [0052]
- GRABOWSKI GATTHOROWITZ Crit Rev Biochem Mol Biol, 1990, vol. 25, 6385-414 [0053]
- HENRISSAT B et al. Proc Natl Acad Sci U S A, 1995, vol. 92, 157090-4 [0053]
- JORDAN et al. Applied Microbiol Biotechnol, 2010, vol. 86, 1647- [0053]
- CANALS et al. Act Crystallogr. D Biol., 2003, vol. 59, 1447-53 [0089]
- GOULDBiotech, and Bioengr., 1984, vol. 26, 46-52 [0138]
- TEIXEIRA et al. Appl. Biochem. and Biotech., 1999, vol. 77-79, 19-34 [0139]
- KNOWLES et al. Trends in Biotechnology, 1987, vol. 5, 9255-261 [0160]
- SHULEIN Methods in Enzymology, 1988, vol. 160, 234-242 [0160]
- KAWAGUCHI et al. Gene, 1996, vol. 173, 287-288 [0165]
- IWASHITA et al. Appl. Environ. Microbiol., 1999, vol. 65, 5546-5553 [0165]
- WONG et al. Gene, 1998, vol. 207, 79-86 [0165]
- MACHIDA et al. Appl. Environ. Microbiol., 1988, vol. 54, 3147-3155 [0165]
- WOOD et al. Nature, 2002, vol. 415, 871-880 [0165]
- CHEN et al. Biochimica et Biophysica Acta, 1992, vol. 121, 54-60 [0169]
- PENTTILA et al. Gene, 1986, vol. 63, 103-112 [0171]
- SALOHEIMO et al. Gene, 1988, vol. 63, 11-21 [0171]
- KVESITADAZE et al. Applied Biochem. Biotech., 1995, vol. 50, 137-143 [0172]
- OKADA et al. Appl. Environ. Microbiol., 1988, vol. 64, 555-563 [0172]
- SALOHEIMO et al. Eur. J. Biochem., 1997, vol. 249, 584-591 [0172]
- SALOHEIMO et al. Molecular Microbiology, 1994, vol. 13, 219-228 [0172]
- OOI et al. Nucleic Acid Res., 1990, vol. 18, 5884- [0172]
- SAKAMOTO et al. Curr. Genet., 1995, vol. 27, 435-439 [0172]
- SAARILAHTI et al. Gene, 1990, vol. 90, 9-14 [0172]
- SHOEMAKER et al. Bio/Technology, 1983, vol. 1, 691-696 [0174]
- TEERI et al. Bio/Technology, 1983, vol. 1, 696-699 [0174]
- SHEIR-NEISS G et al. Appl. Microbiol. Biotechnology, 1984, vol. 20, 46-53 [0178]
- GHOSE, T. K. Pure & Appl. Chem., 1987, vol. 59, 257-268 [0184]
- LUTHI et al. Appl. Environ. Microbiol., 1990, vol. 56, 92677-2683 [0187]
- WINTERHALTER LIEBEL Appl. Environ. Microbiol., 1995, vol. 61, 51810-1815 [0187]
- SIMPSON et al. Biochem. J., 1991, vol. 277, 413-417 [0187]
- KINOSHITA et al. Journal of Fermentation and Bioengineering, 1995, vol. 79, 5422-428 [0187]
- SHARECK et al. Gene, 1991, vol. 107, 75-82 [0187]
- MOROSOLI et al. Biochem. J., 1986, vol. 239, 587-592 [0187]

- KLUEPFEL et al. *Biochem. J.*, 1990, vol. 287, 45-50 [\[0187\]](#)
- BERNIER et al. *Gene*, 1983, vol. 26, 159-65 [\[0187\]](#)
- CLARKE et al. *FEMS Microbiology Letters*, 1996, vol. 139, 27-35 [\[0187\]](#)
- GILBERT et al. *Journal of General Microbiology*, 1988, vol. 134, 3239-3247 [\[0187\]](#)
- DOMINGUEZ et al. *Nature Structural Biology*, 1995, vol. 2, 569-576 [\[0187\]](#)
- NUYENS et al. *Applied Microbiology and Biotechnology*, 2001, vol. 56, 431-434 [\[0187\]](#)
- YANG et al. *Nucleic Acids Res.*, 1998, vol. 16, 14B7187- [\[0187\]](#)
- ZAPPE et al. *Nucleic Acids Res.*, 1990, vol. 18, 82179- [\[0187\]](#)
- ROSE et al. *J. Mol. Biol.*, 1987, vol. 194, 4755-756 [\[0187\]](#)
- REEN et al. *Biochem Biophys Res Commun.*, 2003, vol. 305, 3579-85 [\[0190\]](#)
- SHALLOM et al. *Biochemistry*, 2005, vol. 44, 387-397 [\[0190\]](#)
- ZANOELO et al. *J. Ind. Microbiol. Biotechnol.*, 2004, vol. 31, 170-176 [\[0190\]](#)
- SCHMIDT *Methods Enzymol.*, 1998, vol. 160, 662-671 [\[0190\]](#)
- KURAKAKE et al. *Biochim. Biophys. Acta*, 2005, vol. 1726, 272-279 [\[0190\]](#)
- ANDRADE et al. *Process Biochem.*, 2004, vol. 39, 1931-1938 [\[0190\]](#)
- PINPHANICHAKARN et al. *World J. Microbiol. Biotechnol.*, 2004, vol. 20, 727-733 [\[0190\]](#)
- XUESHAO *Biotechnol. Lett.*, 2004, vol. 26, 1511-1515 [\[0190\]](#)
- KIM et al. *J. Microbiol. Biotechnol.*, 2004, vol. 14, 643-645 [\[0190\]](#)
- OGUNTIMEINREILLY *Biotechnol. Bioeng.*, 1980, vol. 22, 1143-1154 [\[0190\]](#)
- MATSUO et al. *Agric. Biol. Chem.*, 1987, vol. 51, 2367-2379 [\[0190\]](#)
- NUMANBHOSLEJ. *Ind. Microbiol. Biotechnol.*, 2006, vol. 33, 247-260 [\[0193\]](#) [\[0193\]](#) [\[0193\]](#) [\[0193\]](#) [\[0193\]](#)
- OSHIMA et al. *J. Appl. Glycosci.*, 2005, vol. 52, 261-265 [\[0193\]](#)
- KIM et al. *J. Microbiol. Biotechnol.*, 2004, vol. 14, 474-482 [\[0193\]](#)
- SHIN et al. *Appl. Environ. Microbiol.*, 2003, vol. 69, 7116-7123 [\[0193\]](#)
- MARGOLLES et al. *Appl. Environ. Microbiol.*, 2003, vol. 69, 5096-5103 [\[0193\]](#)
- TAYLOR et al. *Biochem. J.*, 2006, vol. 395, 31-37 [\[0193\]](#)
- PANAGIOTOU et al. *Can. J. Microbiol.*, 2003, vol. 49, 639-644 [\[0193\]](#)
- SHALLOM et al. *J. Biol. Chem.*, 2002, vol. 277, 43667-43673 [\[0193\]](#)
- LEE et al. *J. Biol. Chem.*, 2003, vol. 278, 5377-5387 [\[0193\]](#)
- SAKAMOTO et al. *Biophys. Acta*, 2003, vol. 1621, 204-210 [\[0193\]](#)
- RAHMAN et al. *Can. J. Microbiol.*, 2003, vol. 49, 58-64 [\[0193\]](#)
- RAHMAN et al. *Carbohydr. Res.*, 2003, vol. 338, 1469-1475 [\[0193\]](#)
- TUNCERBALL *Folia Microbiol.*, 2003, vol. 48, 168-172 [\[0193\]](#)
- MIYAZAKI *Extremophiles*, 2005, vol. 9, 399-406 [\[0193\]](#)
- JUNG et al. *Agric. Chem. Biotechnol.*, 2005, vol. 48, 7-10 [\[0193\]](#)
- KOSEKI et al. *Biochim. Biophys. Acta*, 2006, vol. 1760, 1458-1464 [\[0193\]](#)
- CHACON-MARTINEZ et al. *Physiol. Mol. Plant Pathol.*, 2004, vol. 64, 201-208 [\[0193\]](#)
- DEBECHE et al. *Protein Eng.*, 2002, vol. 15, 21-28 [\[0193\]](#)
- SORENSEN et al. *Biotechnol. Prog.*, 2007, vol. 23, 100-107 [\[0193\]](#)
- KOTAKE et al. *J. Exp. Bot.*, 2006, vol. 57, 2353-2362 [\[0193\]](#)
- NEOJ. *Wood Chem. Tech.*, 1986, vol. 6, 2147- [\[0200\]](#)
- BEDFORD et al. *Proceedings of the 1st Symposium on Enzymes in Animal Nutrition*, 1993, 73-77 [\[0208\]](#)

- **ERBRINGEROVA et al.** Carbohydrate Polymers, 1998, vol. 37, 231- [\[0224\]](#)
- Determination of structural carbohydrates and lignin in the biomass National Renewable Energy Laboratory 20080000 [\[0226\]](#)
- **PACE et al.** Protein Science, 1995, vol. 4, 2411- [\[0233\]](#)
- **WEICHSELBAUM** Amer. J. Clin. Path., 1960, vol. 16, 40- [\[0234\]](#)
- **GORNALL et al.** J. Biol. Chem., 1949, vol. 177, 752- [\[0234\]](#)
- **SADER et al.** Archives of Veterinary Science, 2004, vol. 9, 273-79 [\[0235\]](#)
- Appl. Biochem. Biotechnol., vol. 161, 313-317 [\[0244\]](#)
- **WALSETHTAPPI**, 1971, vol. 35, 228- [\[0244\]](#)
- **WOOD** Biochem. J., 1971, vol. 121, 353-362 [\[0244\]](#)
- **FUCHS et al.** Fungal Genetics and Biology, 2004, vol. 41, 852-863 [\[0246\]](#)
- **LI et al.** Applied Biochemistry and Biotechnology, 2005, vol. 121-124, 321-334 [\[0246\]](#)
- **SHEIR-NEISS, G et al.** Appl. Microbiol. Biotechnol., 1984, vol. 20, 46-53 [\[0322\]](#)
- **PENTTILA, M et al.** Gene, 1987, vol. 61, 2155-64 [\[0322\]](#)

HIDTIL UKENDTE GLYCOSYLHYDROLASEENZYMER OG ANVENDELSER HERAF
PATENTKRAV

1 Sammensætning, der omfatter:

5 (a) et første polypeptid, hvor polypeptidet har mindst 90 % sekvensidentitet med aminosyresekvensen ifølge SEQ ID NO: 2, eller med aminosyresekvensen, der svarer til resterne (i) 24-766, (ii) 73-321, (iii) 73-394, (iv) 395-622, (v) 24-622 eller (iv) 73-622 ifølge SEQ ID NO: 2, der har β -xylosidaseaktivitet; og

(b) et andet polypeptid, der har xylanaseaktivitet, hvor det andet polypeptid er en *Trichoderma*-xylanase eller en *Aspergillus*-xylanase.

10 2. Sammensætning ifølge krav 1, hvor sammensætningen er en gæringsbouillon, eventuelt cellefri, eventuelt hvor gæringsbouillon er gæringsbouillon af en filamentøs svamp, såsom:

(i) en *Trichoderma*, *Humicola*, *Fusarium*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, *Phanerochaete* eller *Chrysosporium*; eller

15 (ii) en *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 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*,
25 *Penicillium canescens*, *Penicillium solitum*, *Penicillium funiculosum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* eller *Trichoderma viride*.

30 3. Sammensætning ifølge krav 1, hvor sammensætningen er en gæringsbouillon, eventuelt cellefri, eventuelt hvor gæringsbouillon er gæringsbouillon af en filamentøs svamp, såsom:

(i) en *Trichoderma*, *Humicola*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, *Phanerochaete* eller *Chrysosporium*; eller

(ii) en *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium lucknowense*,
35 *Bjerkandera adusta*, *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*,

-2-

Phlebia radiata, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* eller *Trichoderma viride*.

4. Sammensætning ifølge et hvilket som helst af kravene 1 til 3, der endvidere omfatter et tredje
5 polypeptid med en L- α -arabinofuranosidaseaktivitet eller β -xylosidaseaktivitet.

5. Sammensætning ifølge krav 4, hvor det tredje polypeptid er et Fv51A, Af43A, Fv43B, Pa51A, Pf51A, Fv43D, Pa51A, Gz43A, *Trichoderma reesei* Bx11, Pf43A, Fv43E, Fv39A, Fo43A eller Fv43B polypeptid, og hvor:

(a) hvis det er til stede, Fv43D-polypeptidet har mindst 90 % sekvensidentitet med SEQ ID
10 NO: 28, eller med resterne (i) 20-341, (ii) 21-350, (iii) 107-341 eller (iv) 107-350 ifølge SEQ ID NO: 28;

(b) hvis det er til stede, Pf43A-polypeptidet har mindst 90 % sekvensidentitet med aminosyresekvensen ifølge SEQ ID NO: 4 eller med resterne (i) 21-445, (ii) 21-301, (iii) 21-323, (iv) 21-444, (v) 302-444; (vi) 302-445, (vii) 324-444 eller (viii) 324-445 ifølge SEQ ID NO: 4;

(c) hvis det er til stede, Fv43E-polypeptidet har mindst 90 % sekvensidentitet med
15 aminosekvensen ifølge SEQ ID NO: 6, eller med resterne (i) 19-530, (ii) 29-530, (iii) 19-300 eller (iv) 29-300 ifølge SEQ ID NO: 6;

(d) hvis det er til stede, Fv39A-polypeptidet har mindst 90 % sekvensidentitet med aminosekvensen ifølge SEQ ID NO: 8, eller med resterne (i) 20-439, (ii) 20-291, (iii) 145-291 eller 145-439 ifølge SEQ ID NO: 8;

20 (e) hvis det er til stede, Fv43B-polypeptidet har mindst 90 % sekvensidentitet med aminosyresekvensen ifølge SEQ ID NO: 12, eller med resterne (i) 17-574, (ii) 27-574, (iii) 17-303 eller (iv) 27-303 ifølge SEQ ID NO: 12;

(f) hvis det er til stede, Pa51A-polypeptidet har mindst 90 % sekvensidentitet med aminosekvensen ifølge SEQ ID NO: 14, eller med resterne (i) 21-676, (ii) 21-652, (iii) 469-652 eller (iv)
25 469-676 ifølge SEQ ID NO: 14;

(g) hvis det er til stede, Fo43A-polypeptidet har mindst 90 % sekvensidentitet med aminosekvensen ifølge SEQ ID NO: 18, eller med resterne (i) 21-341, (ii) 107-341, (iii) 21-348 eller (iv) 107-348 ifølge SEQ ID NO: 18;

(h) hvis det er til stede, Gz43A-polypeptidet har mindst 90 % sekvensidentitet med
30 aminosyresekvensen ifølge SEQ ID NO: 16, eller med resterne (i) 19-340, (ii) 53-340, (iii) 19-383 eller (iv) 53-383 ifølge SEQ ID NO: 16;

(i) hvis det er til stede, *Trichoderma reesei* Bx11-polypeptidet har mindst 90 % sekvensidentitet med aminosekvensen ifølge SEQ ID NO: 44;

(j) hvis det er til stede, Fv51A-polypeptidet har mindst 90 % sekvensidentitet med
35 aminosyresekvensen ifølge SEQ ID NO: 32, eller med resterne (i) 21-660, (ii) 21-645, (iii) 450-645 eller (iv) 450-660 ifølge SEQ ID NO: 32;

(k) hvis det er til stede, Af43A-polypeptidet har mindst 90 % sekvensidentitet med SEQ ID NO: 20, eller med resterne (i) 15-558 eller (ii) 15-295 ifølge SEQ ID NO: 20;

-3-

(l) hvis det er til stede, Pf51A-polypeptidet har mindst 90 % sekvensidentitet med aminosyresekvensen ifølge SEQ ID NO: 22, eller med resterne (i) 21-632, (ii) 461-632, (iii) 21-642 eller (iv) 461-642 ifølge SEQ ID NO: 22.

6. Sammensætning ifølge et hvilket som helst af kravene 1 til 5, hvor det andet polypeptid er:

5 (i) et *Trichoderma reesei* Xyn3-polypeptid med mindst 90 % sekvensidentitet med SEQ ID NO: 42, eller med resterne 17-347 ifølge SEQ ID NO: 42, eller er kodet for af en nukleinsyre med mindst 90 % sekvensidentitet med SEQ ID NO: 41, eller af en nukleinsyre, der er i stand til at hybridisere under høje stringente betingelser med SEQ ID NO: 41, eller til et fragment deraf;

(ii) et *Trichoderma reesei* Xyn2-polypeptid med mindst 90 % sekvensidentitet med SEQ ID
10 NO: 43, eller med resterne 33-222 ifølge SEQ ID NO: 43;

(iii) et AfuXyn2-polypeptid med mindst 90 % sekvensidentitet med SEQ ID NO: 24, eller med resterne 19-228 ifølge SEQ ID NO: 24, eller er kodet for af en nukleinsyre med mindst 90 % sekvensidentitet med SEQ ID NO: 23, eller af en nukleinsyre, der er i stand til at hybridisere under høj stringens til SEQ ID NO: 23, eller til et fragment deraf; eller

15 (iv) et AfuXyn5-polypeptid med mindst 90 % sekvensidentitet med SEQ ID NO: 26, eller med resterne 20-313 ifølge SEQ ID NO: 26, eller er kodet for af en nukleinsyre med mindst 90 % sekvensidentitet med SEQ ID NO: 25, eller af en nukleinsyre der er i stand til at hybridisere under stringente betingelser til et komplement ifølge SEQ ID NO: 25, eller til et fragment deraf.

7. Sammensætning ifølge et hvilket som helst af kravene 1-6, der endvidere omfatter én eller flere
20 polypeptider med cellulaseaktivitet,

eventuelt hvor det ene eller flere polypeptider med cellulaseaktivitet er uafhængigt komponenter af en fuld cellulase, såsom en β -glucosidase-beriget fuld cellulase, der kan være en filamentøs fuld svampecellulase, der for eksempel er fremstillet af enten (a) *Acremonium*, *Aspergillus*, *Emericella*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Scytalidium*, *Thielavia*,
25 *Chrysosporium*, *Phanerochaete*, *Tolypocladium* eller *Trichoderma species*; eller (b) *Chrysosporium lucknowense*, *Phanerochaete chrysosporium*, *Trichoderma reesei*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma viride*, eller *Pencillium funiculosum*.

8. Sammensætning ifølge et hvilket som helst af kravene 1-6, der endvidere omfatter én eller flere polypeptider med cellulaseaktivitet, eventuelt hvor det ene eller flere polypeptider med cellulaseaktivitet er
30 udvalgt fra en β -glucosidase, en endoglucanase og en cellobiohydrolase, for eksempel et *Trichoderma reesei* Bgl1-polypeptid med mindst 90 % sekvensidentitet med SEQ ID NO: 45, eller med resterne 32-744 ifølge SEQ ID NO: 45,

endvidere eventuelt hvor

(a) β -glucosidasepolypeptidet udgør op til 50 vægt-% eller 2 vægt-% til 50 vægt-% af den
35 totale vægt af cellulaseenzymmer i sammensætningen; og/eller

(b) det ene eller flere polypeptider med cellulaseaktivitet udgør 30 vægt-% til 80 vægt-% af den totale vægt af proteiner i sammensætningen, og, eventuelt, sammen er i stand til at opnå mindst 0,00005 fraktion produkt pr. mg protein pr. gram phosphorsyrebehandlet cellulose (PASC) bestemt ved et calcofluor-assay.

-4-

9. Sammensætning ifølge et hvilket som helst af kravene 1-8, der endvidere omfatter ét eller flere accessoriske proteiner, der er udvalgt fra gruppen bestående af: mannanaser, galactanaser, arabinaser, ligninaser, amyloaser, glucuronidaser, proteaser, esteraser, lipaser, xyloglucanaser, glycosidhydrolase-familie-61-polypeptider, CIP1, CIP2, swollenin, expansiner, CIP1-lignende proteiner, CIP2-lignende proteiner, cellobiosedehydrogenaser og manganperoxidaser, eventuelt hvor:
- 5 (i) det ene eller flere accessoriske proteiner udgør 1 vægt-% til 50 vægt-% af den totale vægt af proteiner i sammensætningen;
- (ii) biomassen er *Miscanthus*, majscolber, majsblade og -stilke, majsfibre, hirse, durra, papir, pulp eller sukkerrørsbagasse.
- 10 10. Isoleret, syntetisk, eller rekombinant nukleinsyre, der koder for et polypeptid med mindst 90 % sekvensidentitet med aminosyresekvensen ifølge SEQ ID NO: 2, eller med aminosyresekvensen, der svarer til resterne (i) 24-766, (ii) 73-321, (iii) 73-394, (iv) 395-622, (v) 24-622 eller (vi) 73-622 ifølge SEQ ID NO: 2, der har β -xylosidaseaktivitet.
11. Ekspressionskassette og/eller vektor, der omfatter en nukleinsyre ifølge krav 10.
- 15 12. Rekombinant værtscelle, der omfatter ekspressionskassetten og/eller vektoren ifølge krav 11, hvor cellen er:
- (a) en celle af en filamentøs svamp, for eksempel en celle af en *Trichoderma*, *Humicola*, *Aspergillus*, *Neurospora*, *Penicillium*, *Cephalosporium*, *Achlya*, *Podospora*, *Endothia*, *Mucor*, *Cochliobolus*, *Pyricularia*, *Phanerochaete* eller *Chrysosporium*, eller
- 20 en celle af en *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *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*
- 25 *miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Neurospora intermedia*, *Penicillium purpurogenum*, *Penicillium canescens*, *Penicillium solitum*, *Penicillium funiculosum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* eller *Trichoderma viride*; eller
- 30 (b) en celle af en bakterie, for eksempel en celle af *Streptomyces*, en *Thermomonospora*, en *Bacillus* eller en *Cellulomonas*.
13. Fremgangsmåde til fremstilling af et rekombinant polypeptid, hvilken fremgangsmåde omfatter: (a) dyrkning af en værtscelle ifølge krav 12; og (b) genvinding af gæringsbouillon, der omfatter polypeptidet.
- 35 14. Saccharificeringsfremgangsmåde, der omfatter behandling af et materiale, der omfatter hemicellulose med sammensætningen ifølge et hvilket som helst af kravene 1-9, eventuelt hvor:
- (i) materialet, der omfatter hemicellulose, er *Miscanthus*, majscolber, majsblade og -stilke, majsfiber, hirse, durra, papir, pulp eller sukkerrørsbagasse;

-5-

(ii) fremgangsmåden giver mindst 60 % xylose fra materialets hemicellulosexylan, der omfatter hemicellulose; og/eller

(iii) fremgangsmåden endvidere omfatter genvinding af monosaccharider.

15. Fremgangsmåde til fremstilling af et biobrændsel, hvilken fremgangsmåde omfatter:

5 (i) etablering af kontakt mellem en sammensætning, der omfatter et biomassemateriale, og enzymsammensætningen ifølge et hvilket som helst af kravene 1-9, til fremstilling af saccharificeret biomasse; og

(ii) etablering af kontakt mellem den saccharificerede biomasse og en gæringsmikroorganisme for at fremstille et biobrændsel.

10 16. Fremgangsmåde ifølge krav 15, hvor biobrændslet er bioethanol.

DRAWINGS

SEQ ID NO:	Nucleotide or Amino Acid	Description
1.	Nucleotide	Nucleotide sequence for Fv3A, a GH3 family enzyme from <i>Fusarium verticillioides</i>
2.	Amino acid	Protein sequence of Fv3A
3.	Nucleotide	Nucleotide sequence for Pf43A, a GH43 family enzyme from <i>Penicillium funiculosum</i>
4.	Amino acid	Protein sequence of Pf43A
5.	Nucleotide	Nucleotide sequence for Fv43E, a GH43 family enzyme from <i>Fusarium verticillioides</i>
6.	Amino acid	Protein sequence of Fv43E
7.	Nucleotide	Nucleotide sequence for Fv39A, a GH39 family enzyme from <i>Fusarium verticillioides</i>
8.	Amino acid	Protein sequence of Fv39A
9.	Nucleotide	Nucleotide sequence for Fv43A, a GH43 family enzyme from <i>Fusarium verticillioides</i>
10.	Amino acid	Protein sequence of Fv43A
11.	Nucleotide	Nucleotide sequence for Fv43B, a GH43 family enzyme from <i>Fusarium verticillioides</i>
12.	Amino acid	Protein sequence of Fv43B
13.	Nucleotide	Nucleotide sequence for Pa51A, a GH51 family enzyme from <i>Podospira anserina</i>
14.	Amino acid	Protein sequence of Pa51A
15.	Nucleotide	Nucleotide sequence for Gz43A, a GH43 family enzyme from <i>Gibberella zeae</i>
16.	Amino acid	Protein sequence of Gz43A
17.	Nucleotide	Nucleotide sequence for Fo43A, a GH43 family enzyme from <i>Fusarium oxysporum</i>
18.	Amino acid	Protein sequence of Fo43A
19.	Nucleotide	Nucleotide sequence for Af43A, a GH43 family enzyme from <i>Aspergillus fumigatus</i>
20.	Amino acid	Protein sequence of Af43A
21.	Nucleotide	Nucleotide sequence for Pf51A, a GH51 family enzyme from <i>Penicillium funiculosum</i>
22.	Amino acid	Protein sequence of Pf51A
23.	Nucleotide	Nucleotide sequence for AfuXyn2, a GH11 family enzyme from <i>Aspergillus fumigatus</i>
24.	Amino acid	Protein sequence of AfuXyn2
25.	Nucleotide	Nucleotide sequence for AfuXyn5, a GH11 family enzyme from <i>Aspergillus fumigatus</i>

Table 1A-1

SEQ ID NO:	Nucleotide or Amino Acid	Description
26.	Amino acid	Protein sequence of AfuXyn5
27.	Nucleotide	Nucleotide sequence for Fv43D, a GH43 family enzyme from <i>Fusarium verticillioides</i>
28.	Amino acid	Protein sequence of Fv43D
29.	Nucleotide	Nucleotide sequence for Pf43B, a GH43 family enzyme from <i>Penicillium funiculosum</i>
30.	Amino acid	Protein sequence of Pf43B
31.	Nucleotide	Nucleotide sequence for Fv51A, a GH51 family enzyme <i>Fusarium verticillioides</i>
32.	Amino acid	Protein sequence of Fv51A
33.	Nucleotide	Nucleotide sequence for Cg51B, a GH51 family enzyme from <i>Chaetomium globosum</i>
34.	Amino acid	Protein sequence of Cg51B
35.	Nucleotide	Nucleotide sequence for Fv43C, a GH43 family enzyme from <i>Fusarium verticillioides</i>
36.	Amino acid	Fv43C protein sequence
37.	Nucleotide	Nucleotide sequence for Fv30A, a GH30 family enzyme from <i>Fusarium verticillioides</i>
38.	Amino acid	Fv30A protein sequence
39.	Nucleotide	Nucleotide sequence for Fv43F, a GH43 family enzyme from <i>Fusarium verticillioides</i>
40.	Amino acid	Fv43F protein sequence
41.	Nucleotide	Nucleotide sequence for Xyn3, a GH10 family xylanase from <i>Trichoderma reesei</i>
42.	Amino acid	Xyn3 protein sequence
43.	Amino acid	Protein sequence of Xyn2, a GH11 xylanase from <i>Trichoderma reesei</i>
44.	Amino acid	Protein sequence of Bxl1, a GH3 β -xylosidase from <i>Trichoderma reesei</i>
45.	Amino acid	Protein sequence of Bgl1, a GH3 β -glucosidase from <i>Trichoderma reesei</i>
46.	Nucleotide	Deduced cDNA for Pa51A.
47.	Nucleotide	Codon optimized cDNA for Pa51A.
48.	Nucleotide	Coding sequence for CBH1 signal sequence upstream of genomic DNA encoding mature Gz43A.
49.	Nucleotide	Coding sequence for CBH1 signal sequence upstream of genomic DNA encoding mature Fo43A.
50.	Nucleotide	Codon optimized nucleotide sequence for CBH1 signal sequence upstream of codon optimized DNA encoding mature Pf51A

Table 1A-2

Enzyme	Accession No.	Strain of Origin
Mg51A	XP_369655 (NCBI)	<i>Magnaporthe grisea</i>
At51A	XP_001212060 (NCBI)	<i>Aspergillus terreus</i>
Pt51A	XP_001939835 (NCBI)	<i>Pyrenophora tritici-repentis</i>
Ss51A	XP_001595513 (NCBI)	<i>Sclerotinia sclerotiorum</i>
Vd51A	EEU33949 (NCBI)	<i>Verticillium dahliae</i>
AfuXyn3	Q4WLV2 (Uniprot)	<i>Aspergillus fumigatus</i>
PfuXyn1	Q8J0K5 (Uniprot)	<i>Penicillium funiculosum</i>
SspXyn1	CAA67143 (NCBI)	<i>Streptomyces</i> sp. S38

Table 1B

Expressed Protein	pNPA (xQuad Delete Background)	pNPX (xQuad Delete Background)
Quad-delete host	1.0	1.0
Fv43A	51.2	43.3
Fv43B	7.9	45.8
Fv43C	11.1	0.7
Fv43D	59.4	715.4
Tr Bxl1	52.2	1289.4
Pf51A	0.3	26.0
Fv51A	34.3	210.1
Pa51A	5.0	45.0
Fv39A	9.7	9.9
Fv3A	8.7	2763.7
Pf43A	77.3	23.1
Fv43E	0.7	0.9
Gz43A	9.4	42.9
Fo43A	5.1	61.4

Table 2

Protein	Activity on Birchwood Xylan at 50°C pH5 (Unit/mg Protein)
AfuXyn2	547
AfuXyn3	1192
AfuXyn5	1092
PfuXyn1	98
SspXyn1	454
Tr Xyn2	554

Table 3

Added Enzymes	Glucose	Xylose	Arabinose
Tr Bxl1	65.3	54.0	4.4
Fv3A	49.8	54.9	17.7
Fv43A	3.4	5.6	2.4
Fv43B	52.6	32.8	15.4
Fv43D	12.0	28.3	4.0
Fv51A	17.4	22.6	10.8
Tr Bxl1 + Fv43A	68.3	63.2	15.2
Tr Bxl1 + Fv43B	66.7	61.5	29.8
Tr Bxl1 + Fv51A	74.4	59.0	15.6
Fv3A + Fv51A	48.4	70.4	70.7
Fv3A + Fv43A	56.9	60.8	16.3
Fv3A + Fv43B	55.4	77.1	64.9
Fv3A + Fv43D	56.6	62.2	15.6
Fv43A + Fv43B	44.0	39.2	43.8
Fv43D + Fv43A	23.6	43.9	19.0
Fv43D + Fv43B	44.8	53.5	23.5
Fv43D + Fv51A	23.5	36.6	14.1
Tr Bxl1 + Fv43A + Fv43B	62.3	73.2	50.4
Tr Bxl1 + Fv43A + Fv51A	59.9	66.0	52.7
Fv43D + Fv3A + Fv51A	56.4	77.1	71.6
Fv43D + Fv43A + Fv51A	17.5	49.3	52.9

Table 4

TABLE 5A

TABLE 5B

Table 5

Table 5A

Run #	Trial #	Ac1000	Quad	Tr Xyn3	Fv43D	Fv51A	Fv43A	Fv43B	Loading ug/mg Carb	Xyl mg/mL	Glu mg/mL	Arab mg/mL	G+X+A mg/mL	Xyl %theor	Glu %theor	Arab %theor
1	40	0.4	0	0.02	0.3	0	0	0.28	21	32.435	25.610	2.632	60.68	70.39	58.94	44.53
2	10	0.4	0	0.152	0.148	0	0	0.3	7	28.847	14.112	2.322	45.28	62.60	32.48	39.29
3	47	0.4	0.3	0.02	0.02	0	0	0.26	21	33.809	29.948	2.706	66.46	73.37	68.93	45.79
4	8	0.66	0	0.02	0.02	0.3	0	0	21	27.911	27.456	2.462	57.83	60.57	63.19	41.66
5	44	0.175	0.066	0.067	0.009	0.3415	0.3415	0	15.99	31.691	16.036	3.226	50.95	68.77	36.91	54.59
6	4	0.4	0	0.3	0.02	0	0	0.28	21	35.440	27.504	2.789	65.73	76.91	63.30	47.19
7	46	0.76	0.2	0.02	0.02	0	0	0	21	26.642	32.461	0.721	58.82	55.65	74.71	12.20
8	37	0.735	0	0.02	0.02	0	0.225	0	6.87	19.295	15.950	0.518	35.76	41.87	36.71	8.76
9	49	0.68	0	0.3	0.02	0	0	0	7	25.472	17.795	0.775	44.04	55.28	40.96	13.11
10	23	0.4	0.3	0.02	0.02	0.26	0	0	21	28.005	28.024	2.415	58.44	60.77	64.50	40.86
11	32	0.4	0	0.129	0.02	0.3	0.067	0.094	21	34.334	25.816	3.167	63.32	74.51	59.42	53.59
12	9	0.4	0	0.02	0.02	0.3	0.26	0	7	22.906	11.955	2.621	37.48	49.71	27.51	44.35
13	29	0.249	0.078	0.0125	0.0125	0	0.461	0.167	33.7	33.305	26.724	3.339	63.37	72.28	61.51	56.50
14	42	0.4	0	0.3	0.02	0.28	0	0	21	32.071	26.609	2.572	61.25	69.60	61.24	43.52
15	11	0.96	0	0.02	0.02	0	0	0	7	20.257	19.160	0.460	39.88	43.96	44.10	7.78
16	43	0.4	0	0.02	0.28	0.3	0	0	7	33.822	14.959	2.673	51.45	73.40	34.43	45.23
17	48	0.753	0	0.227	0.02	0	0	0	21	33.855	34.378	1.257	69.49	73.47	79.12	21.27
18	35	0.4	0.155	0.164	0.158	0	0	0.123	21	36.727	31.042	2.638	70.41	79.70	71.44	44.64
19	1	0.66	0	0.02	0.02	0	0	0.3	21	31.988	28.289	2.644	62.92	69.42	65.11	44.74
20	52	0.68	0	0.02	0.3	0	0	0	7	20.053	15.878	0.375	36.31	43.52	36.54	6.35
21	30	0.4	0.28	0.02	0.3	0	0	0	21	31.742	29.596	0.935	62.32	68.88	68.12	16.67
22	41	0.4	0	0.148	0.3	0.152	0	0	7	29.407	14.066	2.451	45.92	63.82	32.37	41.47
23	21	0.4	0.28	0.3	0.02	0	0	0	21	34.797	32.721	1.456	68.97	75.51	75.31	24.64
24	21	0.68	0	0.02	0.3	0	0	0	21	27.996	28.778	0.825	57.60	60.76	66.23	13.96
25	31	0.4	0.26	0.02	0.02	0	0.3	0	7	24.278	16.366	0.775	41.42	52.69	37.67	13.11
26	34	0.4	0	0.02	0.28	0.3	0	0	21	29.464	23.707	2.538	55.71	63.94	54.56	42.94

27	2	0.4	0	0.02	0.02	0	0.26	0.3	7	24.484	13.033	2.612	40.13	53.13	30.00	44.20
28	12	0.4	0	0.02	0.02	0.26	0.3	0	21	28.054	22.396	2.996	53.45	60.88	51.54	50.69
29	36	0.4	0	0.02	0.02	0.3	0	0.26	7	23.951	13.139	1.672	38.76	51.98	30.24	28.29
30	25	0.4	0.3	0.02	0.28	0	0	0	7	23.772	16.611	0.817	41.20	51.59	38.23	13.82
31	50	0.755	0	0.02	0.02	0.205	0	0	7	22.539	16.844	1.475	40.86	48.91	38.77	24.96
32	19	0.66	0	0.02	0.02	0	0.3	0	21	26.686	27.769	1.083	55.54	57.91	63.91	18.32
33	7	0.4	0.3	0.28	0.02	0	0	0	7	29.285	18.127	1.014	48.43	63.55	41.72	17.16
34	6	0.4	0	0.3	0.3	0	0	0	21	33.763	26.607	1.341	61.71	73.27	61.24	22.69
35	16	0.4	0	0.02	0.28	0	0.3	0	7	20.075	12.639	0.448	33.16	43.57	29.09	7.58
36	26	0.4	0.26	0.02	0.02	0	0	0.3	7	26.736	16.644	1.615	45.00	58.02	38.31	27.33
37	24	0.4	0	0.02	0.3	0	0.28	0	21	25.953	22.638	0.788	49.38	56.32	52.10	13.33
38	22	0.4	0	0.02	0.02	0.26	0	0.3	21	29.727	24.463	2.016	56.21	64.51	56.30	34.11
39	51	0.66	0	0.02	0.02	0	0	0.3	7	23.726	17.116	1.460	42.30	51.49	39.39	24.70
40	38	0.4	0	0.3	0.15	0	0.15	0	21	31.816	26.602	1.490	59.91	69.05	61.22	25.21
41	3	0.552	0	0.02	0.155	0.135	0	0.138	21	30.870	28.460	1.934	61.26	66.99	65.50	32.72
42	14	0.66	0.3	0.02	0.02	0	0	0	7	22.806	19.559	0.714	43.08	49.49	45.01	12.08
43	7	0.4	0	0.02	0.02	0	0.3	0.26	21	31.574	23.883	3.152	58.61	68.52	54.97	53.33
44	17	0.4	0	0.28	0.02	0	0.3	0	7	27.356	14.697	1.005	43.06	59.37	33.83	17.01
45	33	0.4	0.106	0.02	0.3	0	0	0.174	7	26.208	15.846	1.521	43.58	56.88	36.47	25.74
46	45	0.96	0	0.02	0.02	0	0	0	21	24.121	32.328	0.729	57.18	52.35	74.40	12.34
47	20	0.4	0.26	0.02	0.02	0.3	0	0	7	25.246	16.419	1.663	43.33	54.79	37.79	28.14
48	27	0.4	0	0.3	0.02	0	0.112	0.168	7	30.163	14.935	1.941	47.04	65.46	34.37	32.84
49	40	0.4	0	0.02	0.3	0	0	0.28	21	28.688	25.604	1.675	55.97	62.26	58.93	28.34
50	5	0.4	0	0.3	0.3	0	0	0	7	30.932	15.365	1.881	48.18	67.13	35.36	31.83
51	28	0.4	0	0.3	0.02	0.28	0	0	7	27.284	15.235	0.878	43.40	59.21	35.06	14.86
52	15	0.4	0.3	0.02	0.02	0	0.26	0	21	33.731	30.530	1.574	65.84	73.20	70.26	26.63
53	8	0.66	0	0.02	0.02	0.3	0	0	21	30.207	30.286	1.907	62.40	65.55	69.70	32.27
54	10	0.4	0	0.152	0.148	0	0	0.3	7	32.186	15.440	1.878	49.50	69.35	35.54	31.78
55	4	0.4	0	0.3	0.02	0	0	0.28	21	40.441	30.159	2.402	73.00	87.76	69.41	40.64
56	18	0.4	0	0.28	0.02	0	0.3	0	21	35.302	27.592	1.840	64.73	76.61	63.50	31.13

Note: Maximum theoretical xylose, glucose and arabinose were based on 46.08 mg/mL, 53.45 mg/mL, and 5.91 mg/mL, respectively.

Table 5B

			Model Calculated Optimal Proportions						
	Loading Total mg/g Glucose and Xylose	r2 Data Fit to Model (Includes both Loadings)	Fraction Accellerase	Fraction Quad Del Sup	Fraction Purified Xyn3	Fraction Purified Fv43D	Fraction Purified Fv51A	Fraction Purified Fv43A	Fraction Purified Fv43B
Total mg/mL G+X+A	21	0.97	0.4	0.3	0.16	0.02	0	0.04	0.08
	7		0.4	0.3	0.15	0.02	0	0.06	0.07
% Yield Glucose	21	0.99	0.54	0.3	0.14	0.02	0	0	0
	7		0.52	0.3	0.12	0.02	0	0.05	0
% Yield Xylose	21	0.92	0.4	0	0.22	0.02	0	0.1	0.26
	7		0.4	0	0.22	0.02	0	0.11	0.25
% Yield Arabinose	21	0.94	0.4	0	0.13	0.02	0	0.19	0.26

Table 6

Run #	Accellerase® 1000	Xyn3	Fv3A	Fv43D	Fv51A	Fv43A	Fv43B	Glucose Yield	Xylose Yield	Arabinose Yield
20	21	0	0	0	0	0	0	47.1%	23.9%	11.7%
21	21	0	0	0	0	0	0	47.6%	24.0%	11.6%
22	21	0	0	0	0	0	0	46.0%	23.4%	12.2%
1	13.125	7.875	0	0	0	0	0	50.3%	40.1%	17.7%
2	12.474	7.476	0	1.05	0	0	0	55.6%	79.4%	23.1%
3	11.55	7.35	0	2.1	0	0	0	49.3%	43.8%	19.1%
4	10.5	6.3	0	4.2	0	0	0	46.0%	45.3%	19.4%
5	10.5	6.3	0	1.05	0	0	3.15	52.8%	78.0%	31.0%
6	10.5	6.3	0	1.05	0	3.15	0	51.3%	80.0%	28.9%
7	10.5	6.3	0	1.05	3.15	0	0	51.3%	80.4%	33.9%
8	10.5	6.3	0	1.05	1.05	1.05	1.05	52.7%	86.0%	40.4%
9	10.5	6.3	0	1.05	1.575	1.575	0	53.2%	87.3%	39.8%
10	10.5	6.3	0	1.05	0	1.575	1.575	51.5%	88.3%	38.9%
11	10.5	6.3	0	1.05	1.575	0	1.575	50.4%	80.4%	33.8%
12	10.5	6.3	3.15	1.05	0	0	0	53.8%	91.6%	34.3%
13	10.5	6.3	1.575	1.05	1.575	0	0	49.3%	90.1%	50.0%
14	10.5	6.3	1.575	1.05	0	1.575	0	50.4%	85.8%	33.4%
15	10.5	6.3	1.575	1.05	0	0	1.575	49.8%	89.4%	45.5%
16	10.5	6.3	1.05	1.05	0	1.05	1.05	50.2%	93.9%	48.3%
17	10.5	6.3	1.05	1.05	1.05	1.05	0	49.0%	85.2%	47.8%
18	10.5	6.3	1.05	1.05	0	1.05	1.05	48.8%	88.1%	46.1%
19	10.5	6.3	0.7875	1.05	0.7875	0.7875	0.7875	50.5%	91.2%	50.7%

Table 7

Fv3A mg/g Xylan	Fv43D mg/g Xylan	Fv51A mg/g Xylan	>dp2 mg/mL	Cellobiose mg/mL	Xylobiose mg/mL	Glucose mg/mL	Xylose mg/mL
0	0	0	12.8	5.3	26.8	50.7	35.6
0.5	0	0	6.3	11.0	14.6	53.1	50.9
1	0	0	4.1	10.7	9.2	52.5	55.1
2	0	0	3.6	11.1	6.6	53.5	60.4
3	0	0	3.8	11.1	6.7	54.5	64.4
0.5	0	0.5	2.5	5.0	15.7	55.5	56.8
1	0	0.5	2.2	4.2	14.6	56.9	64.5
2	0	0.5	1.5	6.2	8.2	57.5	67.4
3	0	0.5	1.0	6.0	7.5	57.7	69.4
0.5	0	1	5.6	5.9	16.2	55.9	59.2
1	0	1	2.2	5.7	11.9	55.9	63.6
2	0	1	1.0	5.0	8.5	55.7	68.4
3	0	1	0.7	5.9	7.1	57.9	68.1
0.5	0	2	4.4	6.2	14.9	56.1	58.7
1	0	2	1.9	5.1	12.8	55.6	64.8
2	0	2	1.5	6.7	8.4	57.3	69.1
3	0	2	1.5	6.0	6.7	58.7	70.5
0	0.5	2	8.8	5.7	4.1	56.2	64.3
0	1	2	9.3	6.1	4.2	56.0	66.8
0	2	2	8.8	5.1	3.9	55.9	66.2
0	3	2	9.0	4.8	4.1	57.2	66.6

Table 8

Mix	Glucose Yield	Xylose Yield	Arabinose Yield
A	1.3%	64.3%	57.7%
B	1.0%	58.5%	44.0%
C	1.4%	64.3%	58.3%

Table 9

Cob % Yields			Switchgrass % Yields			Sorghum % Yields			Bagasse % Yields		
Mix	Glu	Xyl	Ara	Glu	Xyl	Ara	Glu	Xyl	Ara	Glu	Xyl
A	5.6%	73.2%	69.4%	5.8%	76.2%	71.8%	2.6%	84.8%	71.6%	2.1%	92.6%
B	3.2%	62.6%	42.6%	4.8%	72.0%	64.6%	59.8%	83.1%	70.3%	2.0%	95.9%
C	1.8%	73.3%	68.3%	14.3%	70.6%	67.0%	2.5%	98.5%	81.1%	1.8%	94.6%

Table 10

Protein	H3A	39A	69A	A10A	G6A	102	44A	11A	G9A
Fv3A	9.6%	1.5%	9.4%	10.4%	0.0%	0.0%	9.8%	1.9%	3.3%
Fv51A+Fv43D+EndoH	14.8%	38.1%	6.1%	24.3%	14.8%	38.5%	17.0%	41.9%	15.2%
Xyn3	12.6%	8.0%	15.3%	10.2%	12.1%	10.5%	0.2%	11.1%	14.1%
BGL1	7.5%	5.2%	7.6%	7.5%	9.2%	5.1%	9.4%	4.5%	7.7%
CBH1	36.4%	35.2%	46.5%	30.2%	42.8%	32.1%	48.6%	30.5%	43.7%
EGL's	5.6%	4.1%	1.7%	9.0%	11.9%	6.7%	2.4%	1.7%	3.8%
CBH2	9.5%	7.8%	12.7%	8.3%	9.2%	7.2%	12.4%	8.2%	11.9%
Other	4.03%	0.12%	0.73%	0.10%	0.03%	-0.03%	0.27%	0.28%	0.29%

Table 11

Experimental Parameters		Response 1			Response 2		Response 3	
Time (Min)	Temp (°C)	% Solids	% NH3	% Glucan Conversion	% Xylan Conversion	% Total Conversion		
1	90	47.5	14	49.3%	67.4%	57.1%		
2	130	40	21	49.4%	78.3%	61.8%		
3	90	47.5	14	53.5%	74.4%	62.5%		
4	130	52	23	51.7%	78.0%	63.0%		
5	90	47.5	14	53.5%	76.5%	63.4%		
6	90	47.5	14	47.0%	67.5%	55.8%		
7	90	47.5	14	51.0%	71.4%	59.8%		
8	170	47.5	14	50.0%	64.4%	56.2%		
9	90	45.5	28	60.9%	84.0%	70.8%		
10	90	32.5	14	52.4%	74.4%	61.9%		

Table 12

Dry Solids %	Time (Day)	Enzyme (mg TP/g Glucan)	Glucose (g/L)	Xylose (g/L)	Glucan %	Xylan %
15%	2	40	111.6	29.2	89.0%	89.6%
15%	3	20	98.7	27.1	78.7%	83.1%
20%	3	20	90.2	27.7	54.2%	63.8%
20%	4	20	114.2	30.9	68.6%	71.3%
20%	2-day (fed-batch)	20	92.1	27.1	55.3%	62.4%
20%	3-day (fed-batch)	20	114.7	30.1	68.9%	69.4%

Table 13

Sample Name pH / Temp	g/L			% Conversion		
	Cellobiose	Glucose	Xylose	Glucan (Glucose)	Glucan (Cellobiose + Glucose)	Xylan (Xylose)
5.0 / 45	5.83	71.65	16.22	92.13%	100.01%	80.20%
5.0 / 45	5.04	71.18	15.92	91.52%	98.35%	78.72%
4.65 / 47.2	3.71	76.05	16.94	97.79%	102.81%	83.78%
4.65 / 47.2	3.62	76.21	16.81	97.99%	102.89%	83.09%
5.35 / 47.2	6.59	69.89	16.12	89.87%	98.79%	79.68%
5.35 / 47.2	6.67	70.12	15.69	90.16%	99.19%	77.60%
5.0 / 52.5	3.38	77.32	16.79	99.42%	103.99%	83.00%
5.0 / 52.5	3.61	77.43	16.69	99.56%	104.45%	82.54%
5.0 / 52.5	3.63	77.13	16.76	99.17%	104.09%	82.86%
5.0 / 52.5	3.84	77.65	16.85	99.85%	105.05%	83.31%
5.0 / 52.5	3.41	77.14	16.78	99.19%	103.80%	82.96%
4.5 / 52.5	3.21	77.76	16.95	99.99%	104.33%	83.81%
4.5 / 52.5	3.26	76.88	16.79	98.85%	103.26%	83.04%
5.5 / 52.5	6.41	72.05	16.24	92.65%	101.32%	80.32%
5.5 / 52.5	7.01	71.58	16.34	92.05%	101.53%	80.78%
4.65 / 57.8	2.59	71.04	15.62	91.35%	94.85%	77.21%
4.65 / 57.8	2.73	70.66	15.57	90.86%	94.55%	76.99%
5.35 / 57.8	1.72	58.23	12.49	74.87%	77.20%	61.76%
5.35 / 57.8	1.80	58.91	12.40	75.75%	78.19%	61.29%
5.0 / 60	0.13	36.64	7.37	47.12%	47.29%	36.46%
5.0 / 60	0.15	38.21	7.67	49.13%	49.32%	37.93%

Table 14

SEQ ID NO:1**Nucleotide sequence for Fv3A, a GH3 family enzyme from *Fusarium verticillioides***

atgetgetcaatcttcaggtogetgccagcgctttgtcgctttctcttttaggtggattggctgaggetg
 ctacgccatatacccttcggactgtaccaaaaggacctttgagcaagaatggaatctgogatacttcgtt
 atctccagctaaaagagcggtgctctagttgctgctctgaagccgaagagaaggtgggcaatctggtc
 aggtaaaatatccccccccccataatcactatttcggagattggagctgacttaacgcagcaatgcaactg
 gtgcaccaagaatcggacttccaagggtacaactgggtggaacgaagcccttcattggcctcgctggatctcc
 aggtggctcgctttgcgacactcctccctacgacgcggccacatcatctcccatgcctcttctcatggcc
 gctgctttcgacgatgatctgatccacgatatcggcaacgtcgctggccaccgaagcgctggcttcaacta
 acggcggtttggcgcgaggctcgacttctggacacccaacgtcaaccccttttaaagatcctcgctggggtcg
 tggctccgaaaactccaggtgaagatgccccttcattgacgcgggtatgctcgctatatcgctcaggggtctc
 gaaggcgataaggagcaacgacgtattgttgcctacgtgcaagcaactatgctggaaacgactttgaggact
 ggggaggettcacgcgtcacgactttgatgccaaagattactcctcaggacttggctgagtaactacgtcag
 gcccttccaggagtgcaccgcgtgatgcasagggttggttccatcatgtgcgcctacaatgcgcgtgaacggc
 attcccgcatgcgcaaacctcglatctgcaggagacgaltccacagagggcactggaaclggacgcgcgala
 acaactggatcactagtgtgtggcgccatgcaggatatctggcagaatcacaagtatgtcaagaccaa
 cgctgaagggtgccaggttagcttttgagaacggcatggattctagctgcgagtataactactaccagcgat
 gtctccgattcgtacaagcaaggcctcttgactgagaagctcatggatcgcttctgtgaagcgcttttctg
 aagggtctgttcatactggctttctttgaaggtgccaaagcgcaatggaaactcgctcagttttgcggtatgt
 caacaccaaggaagctcaggatcttgcaactcagatctgctgtggagggtgctgttcttcttaagaatgac
 ggcacttttgctctgaagctcaagaagaaggatagtggtgcaatgatcggattctgggccaacgatactt
 ccaagctgcagggtgggttacagtggacgtgctcgttccctccacagcccgctttatgcagctgagaagct
 tggctcttgacaccaacgtggcttggggtccgacactgcagaacagctcatctcatgataactggaccacc
 aatgctgllgctgcggcggaagaagctcgtatllacallctclactllggcggctcllgacgcctclgctgctg
 gcgaggacagagatcgtgagaaccttgactggcctgagagccagctgaccttcttcagaagctctctag
 tctcggaagccactggctgttatccagcttgggtgatcaagtcgatgacaccgctcttttgaagaacaag
 aagattaacagttattctttgggtcaattaccctggctcaggatggcgccactgcagtcattggacctgctca
 ctggacgaaagagtcctgctggcgcactaccgctcagcgaatatccagtaataactgagcagattgg
 catgactgacatggacctcagacctaccaagtcgcttgccaggggagaacttatcgctgggtactcaactcca
 gttcttccctacggctttggcctccactacaccaagttccaagccaagttcaagttccaacaagttgacgt
 ttgacatccagaagcttctcaagggtgcagtgctcaataactccgataacttgcgcgctgccccccatcca
 agttagtgtcaagaacaccggccgcattactcctcgactttgtctctctcggtcttttatcaagagtgaagtt
 ggacctaaagccttaacctctcaagaccttgcggcttatggctcgcttgcattgatgtgcgcgcttcacga
 cgaaggataltclactlgaglggaagctllggataacaltgcgcgagggggagagaaltgglgaltlggllgl
 ttatcctgggaacttacactctgttgcgtggatgagcctacgcaagccaagatccaggttacgctgactgga
 aagaaggctattttggataagtggtgcctcaagaccccaagctctgcgtaa

FIG. 1A**SEQ ID NO:2****Protein sequence of Fv3A**

mllnlqvaasalslsllgclaeaatpytlpdetkgplskngicdtslspakraaalvaaltpeekvgnlv
 snATGAPRIGLPRYNWNNEALEGLAGSPGGRFADTPPYDAATSEFPMPLIMAAAFDDDLIHDIGNVVGTEA
 RAFTNGGWRGVDFWTFNVNPFKDFRWGRGSETPGEDALHVSRYARYIVRGLEGDKQRRIVATCKHYAGN
 DFEDWGGFTRHDFDAKITPQDLAEYYVRPFQECTRDAKVGSI MCAYNAVNGIPACANSYLQETILRGHWN
 WTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMDssceybttsdvdsdykqgllteklmdrsl
 krlfeglvhtgffdgakagwnslsfadvntkeagdlalrsavegAVLLKNDGTLPLKLKKKDSVAMIGFW
 ANDTSKLQGGYSGRAPFLHSPLYAAEKLGLD TNVANGPTLQNSSSHDNWTTNAVAAAKKSDYILYFGGLD
 ASAAGEDRDRENLDWPFESQLTLLQKLSSLGRPLVVIQLGDQVDDTALLKNKKINSILWVNYPGQDGGTAV
 MDLLTGRKSPAGRLPVTQYPSKYTEQIGMTDMDLRPTKSLPGRTYRWYSTFVLPYGFGLHYTKfqakfks
 nkltfd'qkllkgcsaqysdtcalppiqvsvkntgritdsdfvslvfi'ksevgpkpyplktlaaygrlhdv
 apsstkdislewtldniarrgengdlvypgtytllldeptqakiqvlttgkkaildkwpqdpksa

FIG. 1B

SEQ ID NO:3**Nucleotide sequence for Pf43A, a GH43 family enzyme from *Penicillium funiculosum***

atgcttcagcgaatttgcttatattttaccactggctctattgagtgttggagtgaaagccgacaaccct
 ttgtgcagagcatctacacccgtgatccggcacccgatgggtatacaatgacccgcgtttatgtcttcattgga
 ccatgacaacacccggagctacctactacaacatgacagactggcatctgttctcgtcagcagatatggcg
 aattggcaagatcatggcattccaatgagcctggccaatttcacctgggccaacgcgaatgcgtgggccc
 cgcaagtcacccctcgcaacggccaattctacttttatgctcctgtccgacacaaacgatggttctatggc
 tatcgggtgtgggagtgagcagcaccatcacagggtccataccatgatgctatcggcaaacccgctagtagag
 aacaacgagattgatcccaccgtgttcacgcagatgacggtcaggcataacctgtactggggaaatccag
 acctgtggtaagtcacaaattgaaccaagatatgatatcgtacagcgggagccctactcagattccactcac
 caaggctggattttgggtactcgaacggggcaatgctcaacgggccgaccacttttgaagaagctccatgggta
 tacaaacgcaacggcatctactatatcgccatgacagcggattgttgttctgaggatattcgtactcca
 cgggaaccagtgcactgggtccgtggacttatcgaggcgtcatcatgcgcgacccaaggtagcagcttcac
 caatcacgaggggtattatcgacttccagaacaactcctactttttctatcacaaacggcgctcttccggc
 ggaggcggctaccaacgatctgtatgtgtggagcaattcaaatataatgcagatggaaccattccgacga
 tcgaaatgaccaccgcgggtccagctcaaattgggactctcaacccttacgtgcgacaggaagccgaaac
 ggccgcatgggtcttcaggcatcactacggaggtttgtagcgaaggcgggaattgacgtcgggtttatcaac
 aatggcgattacatcaaagttaaaggcgtagctttcggttcaggagcccattctttctcagcgcgggttg
 cttctgcaaatagcggcggcactattgcaatacacctcgggaagcacaactgggtacgctcgtgggcacttg
 tactgtccccagcactggcggttggcagacttggaactacgttacctgttctgtcagtggcgcactctggg
 acccaggatgtgtattttgttttcgggtggtagcgggaacaggataacctgttcaactttgattattggcagt
 tcgcataa

FIG. 2A**SEQ ID NO:4****Protein sequence of Pf43A**

mlqrfavilplallsvqvkadnpfvqsiytadpapmvynndrvyvfmdhdntgatyynmtdwhlfssadma
 nwqdhgipmslanftwananawapqviprngqfyfyapvrhndgsmalgvgvsstitgpyhdaigkplve
 nneidptvfiddgqaylywgnpdlwyvklngdmisysgsptqiplttagfgtrtgnagrpttfeeapwv
 ykrngiyyiayaadccsedirystgtsatgpwtyrgvimptqgssftnhegiidfqnnsyffyhngalpg
 ggggyqrsvcveqfkynadgtiptiemttagpaqigtlnpyvrqEAETAAWSSGITTEVCSEGGIDVGFIN
 NGDYIKVKGVAFGSGAHSFSARVASANSGGTIAIHLGSTTGTLVGTCTVPSTGGWQTWTTVTCSVSGASG
 TQDVYFVFGGSGTGYLFNFDYWQFa

FIG. 2B

SEQ ID NO:5**Nucleotide sequence for Fv43E, a GH43 family enzyme from *Fusarium verticillioides***

atgaagggtatactggctcgtggcgtggggccacttcttttgacgcgcggcaotggctggcttgattggacace
gtgcgcgcaccaccttcaacaatcctatcatctactcagactttccagataacgatgtattcctcggtcc
agataaactactactacttctctgcttccaaacttccacttcagcccaggagcaccogttttgaagtctaaa
gatctgctaaaactgggatctcatcgggccattcaattccccgcctgaactttggcgacggctatgatcttc
ctcctggctcacggttattacogtggaggtacttgggcacatccctcagatacagaaagagcaatggaca
gtggtaactggatcggtgcacaaacttctggcagacctgggtatacactgcctcatcgccggaagggtcca
tggtacaacaagggaacttcoggtgataacaattgctactacgacaatggcatactgatcgatgacgatg
ataccatgtatgtcgtatacgggttccggtgaggtcaaagtatctcaactatctcaggacggattcagcca
ggtcaaattctcaggtagttttcaagaacactgatattgggggtccaagacttggagggttaaccgcatgtac
aagatcaacgggctctactatatactaaacgatagcccaagtggcagtcagacotggattttggaagtcga
aatcaccctggggcccttatgagtctaaaggtcctcgccgacaaagtcaccccgccctatctctgggtggtaa
ctgcgcgcacacagggtagttctcataaagactcccaatgggtggctggtaacttcatgtcattcacttgggoc
tatcctgcgcgcgcgtcttccgggttcttgcaccgattacgtgggggtagcgatgggtttccccattottgtca
aggggtgctaatggcggtatggggatcatcttaaccaacacttctggcagcgatgggtgtgacaaagaattg
gacaaggactgataccttccgcgggaacctcacttgcctcgtcctgggagtggaaccataatecggacgtc
aactccttcaactgtcaacaacggcctgactctccgcactgctagcattacgaaggatatttaaccaggcga
ggaacacgctatctcaccgaactcatgggtgatcatccaacaggaatagtgaagattgatttctctccgat
gaaggacggcgacccgggcccgggtttcagcggttcgagaccaaagtgcatacatcggtattcatcgagat
aacggaaagtccacaatcgctacgaagcatgggatgaatatggatgagtggaacggaacaacaacagacc
tgggacaaaataaaaagccacagctaatgtgccttcttgggaaggaccaagatctggctgagacttcaacttga
taccaacccagcaggaactggcaacactatcttttcttacagttgggatggagtcaagtatgaaacactg
ggteccaacttcaaaactgtacaatgggttgggcattcttttattgcttacogattoggcattcttcaacttgc
ccgagacgggcttttaggaggtctcgatcaaggttgagttctttcacagctgcatag

FIG. 3A**SEQ ID NO:6****Protein sequence of Fv43E**

mkvywlyawatsltpalaglighrrattfnnpiiydsdfpdndvflgpdnyyyfsasnfhfsggapvlk
skdllnwdlighsiprlnfgdgydlppgsryyrggtwasslryrksngqwywigcinfwqtwvytasspegp
wynkgnfgdnncyydngilidddtmyvvygsgevkvsqlsqdgfsqvksqvfvkntdigvqdlegnrm
kinglyyilndspsgsqtwiwskspswgpyeskvladkvtppisggnspggsliktpnggwyfmsftwa
ypagrlpvlapitwgsdgfpilvkangggwgssyptlpgtdgvtknwtrtdtfrgtslapswewnhnpdv
nsftvnnqtlrtasitkdiygartlshrthgdhptgivkidfspmckgdraglsafrdqsayigihrd
ngkftiatkhgmnmdewngtttldlgqikatanvpsgrtkiwlrlqltdnpgtgntifsyswdgvkyetl
gpnfklyngwaffiayrfgifnfaetalggsikvesftaa

FIG. 3B

SEQ ID NO:7**Nucleotide sequence for Fv39A, a GH39 family enzyme from *Fusarium verticillioides***

atgcactacgctaccctcaccacttttgggtgctggctctgaccaccaacgctcgttgcacagcaaggcacag
 caactgtcgacctctccaaaaatcatggaccggcggaaggcccttgggttcaggcttcatatacgggtggcc
 tgacaacggaacaagcgtcgacacctccataccagatttcttggtaactgacatcaaattcaactcaaac
 cgcggcggtggcgcccaaatcccatcactgggttgggcccagagggtggctatgaaggatacctcggcgcgt
 tcaactcaaccttatccaactatcgccaccacgcgcaagtataacgctgactttatcttgggtgcctcatga
 cctctgggggtgcggatggcgggcaggggttcaaactccccgttccctggcgacaatggcaattggactgag
 atggagttattctggaatcagcttgtgtctgacttgaaggctcataatatgctggaagggtcttgtgattg
 atgtttggaatgagcctgatattgatatcttttgggatcgcccggtggtcgcagtttcttgagtattacaa
 tcgcgcgaccaaactacttcgggtgagttactactgatccatacgtattttacagtgagctgactggtcga
 attagaaaaacacttcccaaaactcttctcagtgggcccagccatggcacattctcccattctgtccgatg
 ataaatggcatacctggccttcaatcagtagcgggtaacaagacagtccttgatatttactcctggcatca
 gattggcgcttgggaacgtgagccggacagcactatccccgactttaccaccttgcggggcgcaatatggc
 gttccccgagaagccaattgacgtcaatgagtacgctgcacgcgatgagcaaaatccagccaactccgtct
 actacclclclcaaclagagcglcaalaaccllagaggclcllccgcgcaaacclgggglagcggatcltgaccl
 ccacaacclggatggggaaccllgalllacagcactaccggllacclcgagggggaccllacclacclaalggcl
 gaatggcaggccttacaagtaactatgcggccatggcagggcagagacttctgaccaaagcactcgtcggact
 tgaagtttgatgtcttttggcactaagcaaggccgtaagattaagattatagccggcagaggaccgttca
 agcaaagtataacatcaaaatcagcgggttgggaagtagcaggacttccctaagatgggtacggtaaggtc
 cggacttatcggttcgactgggtctgggcgcaatggaaagggttgacgggcctgttgatttggggggagaaga
 agtatacttattcggccaatacgggtgagcagccctctacttga

FIG. 4A**SEQ ID NO:8****Protein sequence of Fv39A**

mhyatlttlvlalttnvaaqqgtatvdlsknhgpakalgsgfliygwpdngtsvdtstipdfvlvtdikfn
 rgggaqipslgwarggyegylgrfnstlsnyrttrkynadfillphdlwgadggggsnspfpqdnngnwte
 melfwnqlvsdlkahnmleglvidvwnepdidifwdrpwsqfleyynratkllrktltpktllsgpamahs
 pilsddkwhtwlqsvagnktvpdiyswhqigawerepdstipdfttlraqygvpekpidevneyaardeqn
 pansvyyllsqlerhnlrglranwgsgsdlnwmgnliysttgtsegtyypngewqaykyyaamaggrlvt
 kassdlkfdfatkqgrkikiiaqtrtvqakynikisglevaglpkmgvtvkvrtyrfdwagpngkvdgvp
 dlgekkytysantvsspst

FIG. 4B

SEQ ID NO: 9

Nucleotide sequence for Fv43A, a GH43 family enzyme from *Fusarium verticillioides*

atgtggetgacotccccattgctgttgcgcagcaccctctctgggctcactggcggttgctctagcagaca
accccatcgctocaagacatctacaccgcagaccagcaccaatgggtctacaatggccgcggtctadctctt
cacaggccatgacaaocgaaggctctaccgacttcaacatgacagactggcggtctctttctcgctcagcagac
atgggtcaactggcagcaccatgggtgtccccatgagctttaaagaaccttcagctgggccaacagcagagcct
gggctgggtcaagtcggttgccccgaaacggaaagttttacttctatgttctctgtccgtaatgccaaagacggg
tggaatggctattgggtgtcggtgttagtaccacacatccttggggccctacactgatgcccttggaagcca
ttggctcgagaacaatgagatcgacccaactgtctacatcgacactgatggccaggcctatctctactggg
gcaaccctggattgtactacgtcaagctcaaccaagacatgctctctacagtggtagcatcaacaaagt
atcgctcacaacagctggattcggcagccgcgccgaacaacgcgcagcgtctactacttttcgaggaagga
ccgtggctgtacaagcgtggaaatctctactacatgatctaccgcagccaactgctgttccgaggacattc
gctactcaactggacccagcgcgcactggacottggacttacccgcgggtgtcgatgaacaaggcgggtcg
aagcttcaccaaccatcctggcatcatcgactttgagaacaactcgtaacttcttttaaccacaatggcgct
cttgatggagctagcgggttatactcgggtctgtggctgtcgagagcttcaagtatggttcggacgggtctga
tccccgagatcaagatgactacgcaaggcccagcgcagctcaagtctctgaaccocatatgtcaagcagga
ggccgagactatcgccctgggtctgaggggtatcgagactgaggtctgcagcgaagggtgggtctcaacgttgct
ttcatcgacaatgggtgactacatcaagggtcaagggagtcgactttggcagcaccgggtgcaaagacgttca
gcgcccggtgttgcttccaacagcagcgcggaggcaagattgagcttcgacttggttagcaagaccggtaagtt
ggttgggtacctgcacggtaaacgactacgggaaactggcagacttataagactgtggattgccccgctcagt
ggltgctacitggtagagcgaatcattccttctgctcctcaggggtctcgggtctcctctgllcaactlca
actgggtggcagtttagctaa

FIG. 5A

SEQ ID NO: 10

Protein sequence of Fv43A

mwltsp1lfastllqltqvaladnpivqdiytadpamvyngrvylftghdndgstdfnmtdwrlfssad
 mvnwqhghgvpmslktfswansrawagqvvarngkfyfyvpvrnaktggmaigvgvstnilgpytdalqkp
 lvenneidptvyidtdgqaylywgnpgliyyvklngdmlsysgsinkvslttagfgsrpnnaqrpttfeeg
 pwlykrgnlyymiyaanccsedirystgpsatgpwtyrgvvmnkagrsftnhpgiidfennsyffyhnga
 ldggsgytrsvavesfkygsdglipeikmttgppaqlksINPYVKQEAETIAWSEGIETEVCSEGLNVA
 FIDNGDYIKVKGVDGFGSTGAKTFSARVASNSSGGKIELRLGSKTGKLVGTCTVTTTGNWQTYKTVDCPVS
 GATGTSDLFFVETGSGSGSLENFNWWOFs

FIG. 5B

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

atgagcttctcttgggtattgtgcccccttctagcgatgggaagtgccttctgaaaacgaagacggatg
 ttctgacatacaccaaccctgtccttccaggatggcactcggatccatcgtgtatccagaaagatggcct
 ctttctctgcgtcacttcaacattcatctccttcccagggtcttcccggtotatgcctcaagggatctagtc
 aactggcggtctcatcagccatgtctggaaccgcgagaaacagttgcctggcattagctggaagacggcag
 gacagcaacaggggaatgtatgcaccaaccattcgataccacaagggaaacatactacgtcatctgcgaata
 cctggggcggttgagatattattggtgtcatcttcaagaccaccaatccgtgggacgagagtagctggagt
 gacctgttaccttcaagccaaatcacatcgaccocgatctgttctgggatgatgacggaaagggtttatt
 gtgctaaccatggcatcactctgcaggagattgatttggaaactggagagcttagcccgagcttaatat
 ctggaacgggcacaggaggtgtatggcctgaggggtcccatatctacaagcgcgacggttactactatctc
 atgattgcccagggtggaactgccaagaccacgctatcacaaatcgctcgggcccgcgaagatcaccggcc
 cctatgaagcctacaataacaaccaatcttgaccaaccgcgggacatctgagtacttccagactgtcgg
 tcacgggtgatctgttccaagataccaagggcaactgggtgggggtctttgtcttgcactcgcatacacagca
 caggaggtttcaccatggggccgtgaagctgttttgttcaatggcacatggaacaagggcgaatggccca
 agttgcaaccagtagcaggtgcgcatgcctggaaacctcctcccaaagccgacgcgaaaacgttcccgga
 tgggccccttcaacgctgacccagacaactacaacttgaagaagactaagaagatccctcctcactttgtg
 caccatagagtcaccaagagacgggtgccttctctttgtcttccaagggtctgcacatcgtgcctagtcgaa
 acaacggttacccgtagtgtgttgccaggagatgagattgagctatcaggacagcaggtctagctttcat
 cggacgcgcgcaaaactcacactctgttcaaatatagtgttgatctgacttcaagcccaagtcogatgat
 caggaagctggaatcacccgttttccgcacgcagtttcgaccatategatcttggcattgttctgtcttcta
 caaaccaagggcagcaacaagaaatctaagcttgccttccgatccggggccacaggagctcagaatgttcc
 tgcaccgaaggtagtagcgggtcccgatgggtgggagaagggcgtaatcagctacatategaggcagcc
 aacgcgacgcactacaaccttggagcttcgagccacagaggcaagactctcgacatcgcgacagcatcag
 caagtcttgtgagtgagggcacgggttcatttgttggtagtttgccttggaccttatgctacctgcaacgg
 caaaggatctggagtgggaatgtcccaagggaggtgatgtctatgtgacccaatggacttataagcccggtg
 gcacaagagattgatcatggtgtttttgtgaaatcagaattgtag

FIG. 6A**SEQ ID NO:12****Protein sequence of Fv43B**

mrfswwllcpillamqsalpetktdvstytnpvlpgwhsdpsciqkdglflcvststfisfpglpvyasrdlv
 nwrlishvwnrekqlpgiswktaggqqgmyaptiryhkgtyyviceylgvgdiigvifktnpwdessws
 dpvtfkpnhidpdlfwdddgvycathgitlqeidlctgelspelniwngtggvwpegphiykrdgyyyl
 miaeggtaedhaitiararkitgpyeaynnnpiltnrgtseyfqtvgdglfcdtkgnwwglclatrita
 qgvspmgreavlfngtwnkgewpklopvrgrmpgnllpkptrnvpdggpfnadpndynlkktkkipphfv
 hhrvprdgafslsskgllhivpsrnnvtgsvlpgdeielsgqrglafigrrqthtlfkysvdiidfkpkssd
 qeagitvfrtcfidhidlgivriptnqgsnkksklafrfratgaqnvapkvvpvdpdewekgvislhieaa
 nathynlgasshrgtldiatasaslvsngtgsfvgslgpyatengkgsgvecpkkgdvvyv
 tqwtykpvaeidhgvfvksel

FIG. 6B

SEQ ID NO:14
Protein sequence of Pa51A

mihlkoalalalstgcvaidllfvkssggnkkttdimyglmhedinnsgdgglyaelisnrafqgsekfp
snldnwspvggatltlqlakplssalpyssvnpkkgkgtkdkkvgkylanagfwgmdvkrqkytg
sfhvtgeykgdfevslrsaitgetfgkvkvggskkqkwtekefelyvpfkdapnsnntfvvqwdaeagkd
gsldlnliislfpptfkgrknglridlaqtmvelkptflrfpggnmiegnltldtwkwkyetigplkdrpqn
agvweyqgtlgllveymewaddmnllepivgvfaglaldgsvfpesemgwviqqaldeieflltgdakttk
wgavraklghpkpwkvkweignedwlagrpagfesyiynyrfpmmmkafnekypdikiliaspsifdnmti
pagaagdhhpyltpdefverfakfdnlkskdnvtlligeaasthpnnggiawegdlmpwggsvaeaaifli
sterngdkiiigatyapglrsldrwqwsmtwvqhaadpalttrstswyvrilahhiiretlpvdapagkp
nfdplfyvagksesgtgikfaavynstesipvsllkfdglnegavanltvltgpedpygynpftginvvk
ekttfikagkggkftftlpglsvavleladavkggkgkgkgkgn

FIG. 7B

SEQ ID NO:15
Nucleotide sequence for Gz43A, a GH43 family enzyme from *Gibberella zeae*

atqaaqtccaaqttqttatttcccactcctctcttcttccqttqtcaaaagtctttqccacccaaacqacactqtcctcactaqtatq
atqqactgcqqatccttcqqctcatqctctttaacqacaccttqtqctctaccctctcatqacatcgtatqctqqatttqaqaatq
atcctgatggaggccagtaocccatgagagattaccatgtctactctatctgacaaagatctacggttccctgcggtcgatcaccggt
acggccctgtcagtgaggatgtccctgggctctcgcacagatgtgggtcctgacgctgcccacaaagaaaggcaaatactacct
atacttccctgccaaagacaaggatgatctcttcagaatcggcgttgcgtgtctcaccacccccgggacattcgtecccgaca
agagttggatccctcacaetttcagcatcgaccccgccagtttcgtcgatgatgatgacagagcctacttggcatgggtggtatc
atgggtggccagcttcaacgatggcaggataaagaacaaagtacaaacgaatctggcactgagccaggaaacggcacccgtgccttgag
ccctcagattgccaaagctgagcaaggacatgcacactctggcagagagcctcgcgacatgctcattcttgacccccaaagactggca
agccgtcctttctgaggatgaagaccgacgttcttcgaaggaccctggattcacaaggccaacaaagattttactacctcaacctac
tctactggcacacccactatcttgtctatggcacttcaaagacccctatggtccttacacctaccagggcagaattcttgagagcc
agttgatggctggactactcactctatgtagtgcgaagtaccagggtcagtggtggtatcttttatcagcatgccaagacatctggca
aggactatcttcgccagggtaaaggctaaagaagatttgggtacgatatgcaaaaggaaagatctttgacaaaagaagcccttga

FIG. 8A

SEQ ID NO:16

Protein sequence of Gz43A

mkskllfpllsfvgqslatnddcplitsrwtadpsahvfndtlwlypshdidagfendpdgggqyamrdyh
 vysidkiygslpvdhgtalsvedvpwasrqmwapdaahkngkyylyfpakdkddifrigvavsptpggpf
 vpdkswiphtfsidpasfvddddraylawggimggqlqrwqdknkynesgtepgngtaalspqiaklskd
 mhtlaekprdmilildpktgkpllsededrrffegpwihkrnkiiyltystgtthylvyatsktpygpyty
 qgrilepvdgwtthssivkyqgqwwlfyhdaktsgkdylrqvkakkiwydskgkiltkkp

FIG. 8B

SEQ ID NO:17

Nucleotide sequence for Fo43A, a GH43 family enzyme from *Fusarium oxysporum*

atgcagctcaagttttctgtcttcagcattgctgtttctctctgaaccagcaaatgcgctgcgcaagacacta
 atgacattcctcccttgatcaccgacctctgggtccgcagatccctcggtcatgttttcgaaggcaagct
 ctgggtttaccatctcaccgacatcgaagccaatgttggtcaacggcacaggaggcgctcaatacgcctatg
 agggattaccatacctactccatgaagagcatctatggtaaagatcccggttgctgaccacggcgctcgctc
 tctcagtcgatgacgttccctgggcgaagcagcaaatgtgggctcctgacgcagctcataagaacggcaa
 atattatctgtacttccccgcgaaggacaaggatgagatcttcagaattggagttgctgtctcaacaag
 cccagcggctcctttcaaggccgacaagagctggatccctggcacgtacagtatogatcctgctagctaog
 tegacactgataacgaggcctacctcatctggggcggtatctggggcgccagctccaagcctggcagga
 taaaaagaacttttaacgagtcggtggattggagacaaggctgctcctaacggcaccatgccctatctcct
 cagatcgccaagctaagcaaggacatgcacaagatcacggaaacaccccgcatctcgctcattctcgccc
 ccgagacaggccaagcctcttcaggctgaggacaacaagcgacgattcttcgagggcccttggatccacaa
 gcgcggcaagctttactacctcatgtactccaccgggtgatacccacttcccttgcttaogctaacttccaag
 aacatctacggctccttatacctaccggggcgaagattcttgatcctggtgatgggtggaactactcatggaa
 gtattggtgagtataaggggacagtggtggcttttcttctgctgatgggcatacgtctggtaaggattacct
 tcgacaggtgaaggcgaggaagatctggtatgacaagaacggcgaagatcttgcttcaccgctccttag

FIG. 9A

SEQ ID NO:18

Protein sequence of Fo43A

mqlkflssallfsltskcaaqqdndipplitdlwsadpsahvfegklwvypshdieanvvngtggagqyam
 rdyhtysmkxiygdppvvdhgvalsvddvpwakqqmwapdaahkngkyylyfpakdkdeifrigvavsnk
 psqpfkackswipgtysidpasyvdtneayliwggiiwggqlqawqdkknfneswigdkaapngtnalsp
 qiaklskdmhkitetprdlvilapetgkplqaednkrrffegpwihkrnklyylmystgdthflvyatsk
 niygpytyrgkildpvdgwtthgsiveykqgwwlffadahtsgkdylrqvkarkkiwydkngkillhrp

FIG. 9B

SEQ ID NO:19**Nucleotide sequence for Af43A, a GH43 family enzyme from *Aspergillus fumigatus***

atggcagctccaagttttatcctaccccacaggtatccaatcgtataccaatcctctcttccctggttggc
 actccgatcccagctgtgcctacgtagcggagcaagacacotttttctgegtgaogteactttcattgc
 cttccccgggtcttctcttttatgcaagccgagatctgcagaactggaaactggcaagcaatattttcaat
 cggcccagccagatccctgatcttcgctgcacggatggacagcagtcgggtatctatgogcccactctgc
 gctatcatgagggccagttctacttgatcggttctgtacctgggcccgcagaactaagggcttgctgttcac
 ctgctctgatccgtaacgacgatgccgctggagcgatccgctcgaattcgcggtacatggcatcgacccg
 gatattcttctgggatacgaagggacgggtctatgtcacgtccgcgaggaccagatgattaagcagtaca
 cactcgatctgaagacgggggcgattggcccggttgactacctctggaacggcacccggaggagtctggcc
 cgaggcccgcacattttacaagagagacggatactactacctcatgatcgcagagggaggtagccgagctc
 ggccactcgggagaccatggcgcgatctagaacccggacagggtccctgggagccatacccgccacaatccgc
 tcttgctgaacaagggcacctcggagtacttccagactgtgggcccattgctgagcaggtatgggaa
 cggcaactggtgggcgctggcgttgagcacccgatcagggcctgcattggaagaactatcccatgggtcgg
 gagaacggtgctcgccccgcgcttgggagaagggtgagtgccctgtcattcagcctgtgagaggccaaa
 tgcagggggcgtttccaccaccaataagcaggttccctcgcggcgaggcggtatggatcaagcaaccgca
 caaagtggatttccaggcccggatcgaagataccggcgcaacttccagtactggcgatatcccaagacagag
 gattttaccgtctcctctcggggccaccgcaatactcttcggctcaccacctccttttacaacctcaccg
 gaactcgggaacttcaagccggatgatggcctgtcgttgttatgcgcaaacagaccgacaccttggtcac
 gtacactgtggacgtgtcttttgaccccaaggttgccgatgaagaggcgggtgtgactgttttcccttacc
 cagcagcagcacatcgatcttgggtattgtccttctccagacaaccgaggggctgtcgttgtccttccggt
 tccgctggaaggccgcggtaactacgaaggctcctcttccagaagccacgtgctgttcccaagggaatg
 gtgtggacagaccatccggcttgagattcaggccgtgagtgacaccgagtagtctttgoggtgcccgc
 gctcggcaccctgcacagaggcaaatcatcagccgcgccaactcgttgattgtcagtggtgatacgggac
 ggtttactggctcgttgttggcgtgtatgccacgtcgaacgggggtgccggatccacgcccgcataatat
 cagcagatggagatacgaaggacggggccagatgattgattttggtcagtggtcccgagctactga

FIG. 10A**SEQ ID NO:20****Protein sequence of Af43A**

maapslsyptgiqsytnplfpgwhsdpscayvaeqdtffcvststfiafpglplyasrdlqnwklasnifn
 rpsqipdlrvtdgqqsgiyaptlryheggfyilivsylvpqtkgllftssdpyddaawsdplefavhgidp
 difwdhdgtvyvtsaedqmikqytlldlktgaigpvdylwngtggvwpegphiykrdgyyylmiaeggtel
 ghsetmarsrtrtgpwepypnhpnlisnkgstseyfqtvgadlfqdgngnwwavalstrsgpawknypngr
 etvlapaawekgewpvlcpvrqmqgpfpppnkrvprgeggwikpdkvdfprgskipahfqywrypkte
 dftvsprghpntlrtpsfynltgtadfkpddglslvnrkqtdtlftytvdvsfdpkvadecagvtvflt
 qqghidlgivllqtteglslsfrfrvegrgnyecplpeatvpvpkewcgqtirleiqavsdteyvfaaap
 arhpaqrqiisranslivsgdtgrftgslvgvyatsnggagstpaylsrwryegrqgmidfgrvpsy

FIG. 10B

SEQ ID NO:21
Nucleotide sequence for P151A, a GH51 family enzyme from *Penicillium funiculosum*

atgggaaagatgtggcattcgcattcttggttgtgttggtggtctgtctgtcgggcattgccatcactatcacaacgtgtccccaaagtgg
cggcaataagaccagtccttttgcaatacggctctgatgttcgaggtaatccttcttataccacatatataaagtgtcggtcatttct
aagacaagtcaaggacataaaatcacggcgtgatggcgtctgtatgcagagcttgttcgaaacccgagcattccaaaggtagcaccg
tctatccagcaaacctcgatggatcgcactcgggtcaacggagcactccttagcgttccagaatttgacaaaacccctctatcacccctcc
atgcctagctctctcaacgtcgccaagggtccaaacaatggaaagcatcggttttcgcaaatgaaggctgggtggggtatagaagtcaaa
gcccgaagatacgcgggtcattctacgtccagggggaactatcaaggagatcttcgacatctctcttcagtcgaaaaattgacacaag
aagtcttcgcaacggcaaaagtcaggctcctcgggcaaacacagaggactgggttcaatacaagtcaggttggtgccccaaaaaggca
gcatcaaacaccaataaacactctgaccattactcttgactcaaaeggtatgtttaaatcttggttagttcgatgtctgtggcaattgt
cttaacgagaaaacgtlagggattgaaagacggatccttgaaacttcaacttgatcagcctatttcccccaacttacaacaatcggcccc
atggcctaaagaatcgacctggttgaagctatggctgaactagaggggttaagctctttacaaaatcaacttcttcttacgaagacta
atgtgaaaaacttagaaaattctcggcgtctccaggcgttcagcagatgtggaagggtgtacaagctccttactgtgtataaagtggaatgaa
acggtagggagatctcaaggaccggttatagtaggccccagtgcatggacgttaogaagaaagcaatggaaattggcttgattgagtacat
gaattggtgtgatgacatggggcttgagccgagtgagtgattccattcagcgtcaaaatccagtggttctaaatcatcacacatcagtt
cttgccgtatgggatggacattaccttccgaacgaaagtgatatcggaaaaacgatttgcagccatatatcgaagcacacccctcaacca
actgggaattcctgatgggtgccccagatacggccatatggttagttggcgtgcgtctcttggttatccgaagccgtggacgatttaact
acgtcgagatttggaacggaagacaatcctatacgggggaactagaaacatcacatcgccctaccggttttcaggcatatttacgacgctata
acagctaaaatatccccatatgacgggtcatcggaaatctttgacgggagatgccttggtccggcggccgtgcgaagcgattaccatcaata
ttctactcctgatgggtttgtttccccagttcaactactttgatcagatgccagtcaacaaagacactgaacgggtatgaaaaaccc
cccccttttttaaatatgcttttaattggcatttaaaccatctttcatalaggagagatttgaacccgtttatccaaaataatcctagttaattc
ggtggcctggggaagcccatctccccttgatccttggtggattgggtccgttgcaagaagctgttttccctaattggtgaagagagga
attcgccaaaagataatcgggtgctagctacgtacggaaattctacttttcgagatttttaacattgggataagaagggaactaacctcaata
caggctccaatgttccagaaatatcaacaaattgggcagtggtctccaaacactcatcgcttttgacgctgactcgtcgcgtacaaagtgc
ttcaacaagctggcatgtgatcaagggtatcctaaattttcctcctcattcaaaacccgcagatgtgagcctaactttccggaagcttctc
tcgacaaaacaaaatcacgcgcaaaatttacccacgaacttgagtggtgggtgacatagggtccattatactgggttagctggacgaaaacga
caatacagggtcgaacatatattcaaggccgctgttttacaacagcaacctcagacgtcccgtcacccgttcaattttgcagggatgcaacg
caaaagcgcgcaaatttgaccatcttgtcatccgacgacccgaacgcgcatcgaaactacccctggggggccggaagtgtggaagactgag
atccagttctgtcactgcaaaatgctcatggagcatcttgagttcagttctccgaacctaaagtgtggctgttctcaaaaacgggagtaa

FIG. 11A

SEQ ID NO:22

Protein sequence of Pf51A

mqkmwhsilvvlglisvqhaitinvsqsggnktsplqyglmfedinhggdgglyaelvrnrafqgstvyp
 anldgydsvngailalqlntnplspssmpsslnvakgsnngsigfanegwwgievkkparyagsfyvqgdyy
 gdfdislqskltqevfatakvrssgkhedwvqkyelvpkkaasntntltitfdskglkdgslnfnlis
 lfpttynnrpnnglridlveamaelegkflrfpggsdvegvpqpywykwnetvgdlkdrysrsawtyees
 ngiglieymnwcdmglepilavwdghylsnevisendlqpyiddtlncleflmgapdtpygswraslgy
 pkpwtinyveignednlyggletyiayrfgayydaitakyphtvmesltempgpaaaasdyhqystpdg
 fvsqfnfyfdqmpvtntlngeiatvypnnpnsnvawgspfpplypwwigsvaeavfligeernspkiigas
 yapmfrninnwqwsptliafdadssrtrsrtstshvikllstnkitqnlpttwsggdigplywvagrndnt
 gsnifkaavynstsdvpvtvqfagcnaksanltlssddpnasnypggpevkteiqsvtanahgafefs
 lpnlsvavlkte

FIG. 11B

SEQ ID NO:23

Nucleotide sequence for AfuXyn2, a GH11 family enzyme from *Aspergillus fumigatus*

atqgtttctttctcctacctgctgctggcgtgctccgccattgqagctctggtgcccccgctcgaacccg
 agaccacctcgttcaatgagaactgctcttcatgagttcgctgagcgcgcgcggcaccaccaagctccaccgg
 ctggacaacaggctactactactccttctggactgatggcggcgcgacgtgacctacaccaatggcgcc
 ggtggctcgtactccgtcaactggaggaacgtgggcaactttgtcggtggaaaagggtggaaacctggaa
 gcgctaggtaccgagctttgtcaacgtcggatgtgacagacctgtggctgacagaagtagaaccatcaact
 accggaggcagcttcaaccccagcggcaatggctacctggctgtctacggctggaccaccaacctctgat
 tgagtaactacgttgttgagtcgtatggtacatacaaccccggcagcggcggtaccttcaggggcactgtc
 aacaccgacgggtggcacttacaacatctacacggccgcttcgctacaatgctccctccatcgaaggcacca
 agaccttcacccagtaactggtctgtgcgcaacctccaagcgtaccggcggcactgtcaccatggccaacca
 ctccaacgcctggagcagactgggcacgaacctgggaactcacaactaccagattgtcgccactgaggggt
 taaccagagcagcggatctgcttccatcactgtctactag

FIG. 12A

SEQ ID NO:24

Protein sequence of AfuXyn2

mvsfsl_llaacsaidalaapvepettsfnetalhefaeragtpsstgwnngyyysfwtdgggdvtytnga
 ggsysvnwrnvgnfvvgkgwnpgsartinyggsfnpsgngylavygwttmplieyyvvesygtynpgsgg
 tfrgtvntdggtyniytavrynapsiegtktftqywsvrtskrtggtvtmanhfnawsr_lgmnlgtthnyq
 ivategyqssgsasitvy

FIG. 12B

SEQ ID NO:25**Nucleotide sequence for AfuXyn5, a GH11 family enzyme from *Aspergillus fumigatus***

atgatctccattttcctcgcctcagcttttggactcgcgcctatcgcgcgggcgcctatgctcttcgcgagtgaca
aatccgtcagcttagcggaacgtcagacgatcacgaccagccagacagggcacaacaatggctactacta
ttcctttctggaccaacgggtgcgggatcagtgcaatatacaaatgggtgctgggtggcgaatatagtgtgacg
tgggcgaaccagaacgggtgggtgactttacctgtgggaagggctggaatccagggagtgaccagtaggcaa
cgcccgagaactatagaagaggacgcaaagaaagcactaaactctactagtgcattacettctctgg
cagcttcaatecttcoggaaatgcttacctgtccgtgtatggatggactaccaacccctagtggaatac
tacatcctcgagaactatggcagttacaatcctggctcgggcctgacgcacaagggcaccgtcaccagcg
atggatccacctacgacatctatgagcaccaacagggtcaaccagccttcgatcgtcggcaccggccacctt
caaccaatactgggtccatccgcgcaaaaacaagcgatccagcggcacagtcaccacgcgaatcacttcaag
gcctgggctagctctggggatgaacctgggtacccataactatcagattgtttccactgagggatatgaga
gcagcgggtacctcgaccatcactgtctcgtctgggtggttcttctctctggtggaagtgggtggcagctcgtc
tactacttctcaggcagctccctactggtggctcgggcagtgtaagtcttcttccatattggttgtggc
tttatgtgtattctgactgtgatagtgtctgtcttgtggggccagtgcggtggaattggctggctctggt
cctacttgcctcctctcgggcacttgccagggtttcgaaactcgtactactccagtgcttgttagtaccttc
ttgcagggttatatccaagtga

FIG. 13A**SEQ ID NO:26****Protein sequence of AfuXyn5**

MISISSLSFGLAAIAGAYALPSDKSVSLAERQTITTSQTGTNNGYIYSFWTNGAGSVQYTNGAGGEYSVT
WANQNGGDFTCGKGWNPGSDHDITFSGSFNPSGNAYLSVYGWTTNPLVEYYILENYGSYNPGSGMTHKGT
VTSDGSTYDIYEHQQVNQPSIVGTATFNQYWSIRQNKRRSSGTVTTANHFKAWSLGMNLGTHNYQIVSTE
GYESSGTSTITVSSGGSSSGGSSGSSSTTSSGSSPTGGSGSCSALWGQCCGIGWSGPTCCSSGTCQVSNS
YYSQCL

FIG. 13B

SEQ ID NO:27**Nucleotide sequence for Fv43D, GH43 family enzyme from *Fusarium verticillioides***

atgcagctcaagttttctgtcttcagcattgttgctgtctttgacgggcaattgogctgogcaagacacta
 atgatatccctcctctgatcaccgacctctggtctgoggatccctcggtcatgttttcgagggcaaact
 ctgggtttaccatctcacgacatcgaagccaatgtcgtcaacggcacgggagggcgtcagtaacgcatg
 agagattatcacacctattccatgaagaccatctatggaaaagatcccgttatcgaccatggcgctcgctc
 tgtcagtcgatgatgtcccatggggccaagcagcaaatgtgggctcctgacgcagcttacaagaacggcaa
 atattatctctacttccccgccaaggataaagatgagatcttcagaattggagttgctgtctccaacaag
 cccagcgggtcctttcaaggccgacaagagctggatccccggtaacttacagtatcgatcctgctagctatg
 tcgacaactaatggcgaggcataacctcatctggggcggtatctggggcgggccagcttcaggcctggcagga
 tcacaagacctttaatgagtcgtggctcggcgacaaaagctgctcccaacggcaccaacgccttatctct
 cagatcgccaagctaagcaaggacatgcacaagatcaccgagacaccccgcgatctcgtcacccctggccc
 ccgagacagggcaagcccttcaagcagaggacaataagcgacgatttttcgagggggccctgggttcacaa
 gcgcggaagctgtactacctcatgtactctaccggcgacacgcacttccctcgctctacggcacttccaag
 aacatctacggtccttataacctatcagggcaagattctcgacctgttgatgggtggactacgcattggaa
 gtattgttgagtacaagggacagtggtggttgttctttcgggatgcgcatacttctggaaaggattatct
 gagacaggtaaggcgagggaagatctggtatgacaaggatggcaagattttgcttactcgctcctaagatt
 tag

FIG. 14A**SEQ ID NO:28****Protein sequence of Fv43D**

mglkflssalllsltncaaqdndipplitdlwsadpsahvfeqklwvypshdieanvvngtggagqyam
 rdyhtysmkti ygkdpvidhgvalsvddvpwakqqmwapdaaykngkyylyfpakdkdeifrigvavsnk
 psgpfkadkswipgtysidpasyvdtngcayliwgggiwggqlqawgdhktfneswlgdkaapngtnalsp
 qiaklskdmhkitetprdlvilapetgkplqaednkrrffegpwvhkrklyylmystgdthflvyatsk
 niygpytyqgkildpvdgwtthgsiveykgqwwlffadahtsgkdylrqvkarkiw ydkdkilltrpki

FIG. 14B

SEQ ID NO:29

Nucleotide sequence for Pf43B, GH43 family enzyme from *Penicillium funiculosum*

atgagtcgcagcacccttcggtacgcctctgttttcgcctcctggggggggctatgcgcgaaccgtttt
 tggttctcaatagcgattttcccgatcccagtcctcatagagacatccagcggatactatgcattcggtac
 cacoggaaacggagtcgaatgcgcaggttgcttcttcaccagactttaatacctggacttttgccttccggc
 acagatgccctcccgggaccatttcggtcatgggttagcttcgtctccacaaatctggggcgccagatgttt
 tggttaaggtatgttcttatggaataacagtttttaggagtaggtccagccaggatattgacaaaattataa
 taggcgatggtaacctatgtcatgtacttttgggcactctgctgogagtgactcggggcaaacactgcgttg
 gtgcgcgaactgcgacctcacgggaaggacottacaccccggtcgatagcgcctgttgccctgtccattaga
 ccaggaggagctattgatgccaatggattttattgacaccgacggcactatatacgttgtatacaaaatt
 gatggaaacagtctagacgggtgatggaaccacacatcctaccccatcatgcttcaacaaatggaggcag
 acggaaacaaccccaacccggcagcccaatccaactcattgaccgatccgacctcgacggacctttgatcga
 ggctcctagttttgctcctctccaatggaatctactacctcagttttctcttccaactactacaacactaat
 tactacgacacttcatacgcctatgcctcgtcgattactgggtccttggaacaaacaatctgogccttatg
 cacccttggttggttactggaaccgagactagcaatgacgggcgcatgagcgcctcctgggtggtgcgattt
 ctccgtcgatggcaccaagatgttggttccacgcaaacctcaatggacaagatatctcgggcggacgcgc
 ttatttgcctgcgtcaattactgaggccagcgatgtggttacattgcagtag

FIG. 15A

SEQ ID NO:30

Protein sequence of Pf43B

msrsilpyasvfallggaiaepflvlnsdfpdpslietssgyyafgttgngvnaqvasspdfntwtllsg
 tdalpgpfpwvasspqiwapdvlvkadgtyvmyfsasaasdsghcvgaatatspegpytpvdsavacp
 ldqggaidangfidtdgtiyvvykidgnsldgdgtthptpimlqgmeadgttptgspiqldrsdldgpl
 ieapslllsngiyylsfssnyyntnyydtayayassitgpwtkqsapyapllvtgtetsndgalsapgga
 dfsvdgkmlfhanlngqdisggralfaasiteasdvvtlq

FIG. 15B

SEQ ID NO:31
Nucleotide sequence of Fv51A, a GH51 family enzyme from *Fusarium verticillioides*

atggttcgcttcagttcaatcctagcggctgoggcttgcttgctggtcagtcagtcacatcaaggctcgacag
caaggcgggaaaaagctactagcgggtcaccaaatatggcttccttcacgaggttggtattgacacacacacatggcgatga
ttgggatgctaaacttggagctaggatatcaaceaatccgggtgatggtgcatctacgctgagctcatccgcaatcgt
gccttcocagtacagcaagaaataccctgtttctctatctggtggagacccatcaacgatgctaaagctctccctcaa
ccgtctogacactcctctctccgacgctctcccggtttccatgaaagtgaagcctggaaaaggccaaggccaaggaga
ttgggtttcctcaacgagggttactggggaatggatgtcaagaagcaaaaagtacactggctctttctgggttaagggc
gcttacaaaggccacttttacagcttctttgcatctaaaccttaacgacgatgtctttggcagcgtcaagggtcaagtc
caaggccaacaagaagcagtggttgagcatgagtttgtgttacttaacaaagaatgccctaacagcaacaaca
ctttttgctatcacctacgatcccaagtgagtaacaatacaaaactgggaacgtgatgtatatactgacaattttgtagggc
gctgatggagctcttgactcaacctcaattagcttgtttccctcccaacctacaaggccgcaagaacggtctcttcgagt
tgatctttgocgaggctcttcgaagggtctccaccccgtaaggttttaccggtctcaogtgttatcgtgaacagtcgctgact
tgtagaaaaagagccctgctgcgcttccccggtggtlaacatgctogaggggcaacacacaacagacctgggtgggactgga
aggataccctcoggacctctccgcaacccgtcctcgttttcgagggtgtcttgaaactaccagcagacccatggtctcttggga
atctttggagtaacctccagtgggctgaggacatgaaccttqaaatcagtaggttctataaaaattcagtgacggttatg
tgcatgctaacagatttcagttgtcgggtgtctacgctggcctctcctcctcgacgggtcctcgtaaccccccaaggaccaac
tcacgccccctcatcgacgacgctctcgacgagatcgaattcaccgaggtcccgtaactcaaaagtggggaaagaag
cgcgctgagctcggccacccccaaagcctttcagactctcctacgcttgaagtcggaaacgaggaactggctcgcgtgtta
tccccctggctggaaactcttaaaaggagtaccgcttccccatgttccctcgaggctatacaagaaagctacccccgac
tcacgcgtcatctcctctggtgtctctattgacccccgttggttaagaaggatgctgggttcgatattcctgctcctgga
atcgtgactacacaccccttacgcgcgagcctgatgttctgttgaggagttcaacctgtttgataacaataaagtatgg
tcacatcatattgggtgaggttgctttctaccccccaacgggtggaactggctggagtggtaaccttatgaccttaacct
ggtggaatctctggtgtttggcgaggccgctcgtctctgtcggttatgagcgcgaacgcgcgatcgtattccccggaacattc
tacgctcctatcctcaagaacgagaacccgttggcagtggtctatcaccatgatccaattcgcgcgactcgcgcct
gaccacccgctccacacagctgggtatgtctgtggtcaactcttcgcagggccacccccatgacccatactctccccacaccc
ccgacttcgacccccctctactacgtcgtggttaagaacgaggacaagggaactcttattcttggaagggtgctgcgtat
aacacacccaagggtgctgacgttcccggtctctctgtccttcaagggtgtcaagccccgtgtcgaagctgagettac
tctttctgacccaacaaggagaaaggatccttttgcgttcaatgatcctcacaaagggcaacaatgttgttgatactaaga
agactgttctcaaggccgatggaaaagggtgctttcaacttcaagcttccctaacctgagcgtcgtgttcttggagacc
ctcaagaagggaagccttactctagctag

FIG. 16A

FIG. 17A

SEQ ID NO:34

Protein sequence of Cg51B

maplsrlalsllaltgaaaavtlsvansggndtspymyginfedinqsgdgglyaelirnralfhnsslqa
 wtavgdstlevvtsaplsdalprsvkvtsgkgkaglkknagywgmdvqktdkysgsfysygydgkftls1
 vsditnetlattkiksrsvehawtehkfellptksaansnnsfvlefrpchgqtelqfnllslfpptyknr
 pngmrrelmekladlkpsflripggnnlegnyagnywnwsstlgpltdrpggrdgvwtyantdgiglvveym
 hwaedldvevvlavaaglylngdvpeeeelhvfvedalneleflmgdvstpwgarraklgypkpwnikfv
 evgnednlwggldsyksyrktfydaikakypdisifsstdefvykesgqdyhkytrpdysvsqfdlfdn
 wadghpiliigeyatignntgkledtdwdapknkwsnwigsvaeavfilgaerngdrvwgttfapilqnl
 syqwapdlisftanpadttsvsypilqlashrithtlpvssadafgpaywvagraddgsyilkaavy
 nstggadvprvqfeagggggggggggggggdggkkgkkgkkggeggegkvkkgdraqltvltapegpwahnt
 penkgavkttvttkagrggvfefs1pdlsvavlvvegek

FIG. 17B

SEQ ID NO:35

Nucleotide sequence for Fv43C, a GH43 family enzyme from *Fusarium verticillioides*

atgcgtcttctatcggtttccagccatctcctcgtggccttcctaaccctcaaagaggcttcacccctcg
 ccctcagcaaaccgggatagccctgtcctcccgccctctctggggcggaaccccaacatcgccatcgtegacaa
 gacatactacatcttccctaccaccgacgggtttcgaaggctggggggcggaacgtcttctactgggtggaaa
 tcaaaagatctcgtatcctggacaaagagcgacaaagccattccttactctcaatggtagaatggcaacg
 ttccctgggctacaggtaatgcctgggctcctgctttcgctgctcgcgagggaaggtattacttctacca
 tagtgggaataatccctctgtgagtgatgggcataagagtattgggtgcggcggtggctgatcctctgag
 gggccgtggaaggcacaggataagccgatgatcaagggaacttctgatgaggagattgtcagcaaccagg
 ctatcgatcccgctgcctttgaagaccctgagaactggaaagtgggtatatctactggggaaaacgggtgtccc
 cattgtcgcagagctcaacgacgacatgggtctctctcaaaagcaggctggcacaaaatcacagggtcttcag
 aatttccggcaggggtcttttctgtaactatcgcgatggaaacatctctgacatactctatcgacgata
 cgggctcagagaactatcgcggttggttacgctacggcgataaacccattggaccttggacatatcggtgg
 tggttcttctggagaaggacgaatcgaagggcattcttgctacgggacataactccatcatcaacattcct
 ggaacggatgagtggtatatcgcggtatcatcgcttccatattcccgatggaaatgggtataataggggaga
 ctacgattgatagggtaccatcgacaaggatacggggttggttggaagggttacgccgactttgcagag
 tggttgatcctagccctttgtag

FIG. 18A

SEQ ID NO:36

Protein sequence for Fv43C

mrllsfpslllvaflltkeasslalskrdspvlpglwadpniaivdktyyifpttdgfegwggnvfyywk
 skdlvswtksdkpfltlngtngnvpwatgnawapafaarggkyyfyhsgnnpsvsdghksigaavadhpe
 gpwkaqdkpmikgtsdeeivsnqaidpaafedpetgkwyiywngvpivae1nddmvslkagwhkitglq
 nfreqlfvnyrdgtyhltysiddtgseayrvgyatadnpigpwtyrgvlllekdeskgilatghnsiinip
 gtdewyiahyhrfhipdgngynrettidrvpidkdtglfgkvtp1lqsvdprpl

FIG. 18B

SEQ ID NO:37**Nucleotide sequence for Fv30A, a GH30 family enzyme from *Fusarium verticillioides***

atgctcttctctcgctcggttcttctctaccccttgcccttcaagccagccctggcgctcggcgatacatccgtta
 ctgtcgacaccagccagaaaactccaggtcatcgatggcttttggtgtctcagaagcctaaggccacgcca
 acaattccaaaacctcggtcctggaccacagaaagagggcctcgatcttctcttcaacactacaaccggc
 gcaggttatccatcatccgaaaadaagatcggctgcgacgcctccaactccatcaccagcaccaacacccg
 acaacccagataagcaggtctgtttaccattttgaecggcgatgatgatgggtcaggtatggtttagcaaaca
 ggccatgagctatgggtgtagatactatctacgctaatactgttggtctgcgcctgtatacatgaagtcagcc
 cagagtatgggcgctctctgcgggtacacctgggtgtgtcgtgctcctctggagattggagacatcgttacg
 ttgagatgatagctgagtaacctctcctactacaagcaggtcggcatcccagtgctgcacggttggttccct
 caatgaggggtgacggctcggacttttatgctctcaactgcgaacaggtcgcagatgtcattcctcttcta
 cacagcgctttgcagtccaagggccttgggcgatatcaagatgaecgtgctgtgataacatcggttggaagt
 cacagatggactataccgccaagctggctgagcttgaggtggagaagtatctatctgtcatcacatccca
 cgagtaactccagcagccccaaccagcctatgaacactacattgccaacctggatgtccgagggagctgcc
 aatgaccaggcatttgccacagcgtgggtacgtcaacggcggttccaacgaagggtttcacatgggcagtca
 agatcgcacaaggcatcgtaaatgccgacctctcagcgtatatctactgggagggcggttgagaccaacaa
 caaggggtctctatctcactcgtcctcgacacggacggtaaccaagtttaceatatcctcgattctctgggcc
 attgctcactggctcgcgcctatattcgccctgggtgcgcctagactttcgaacttcaggtgttggtgcaagata
 cgattgttggtgcgtttgagaacggttgatggcagtgctcgtcatgggtgctcaccaactctggcactgctgc
 tcagactgtggacctgggtgtttcgggaagtagcttctcaacagctcaggttttcaacttcggatgctgag
 ggcagatggctcgataccaaggtgactctgtccgacggctcgtgtcaagggttacgggtcccggtgcacgggtg
 tcgtcactgtgaagctcacaacagcaaaagctccaaacgggtctcaactgctgtttctgcgcaatctgc
 cccactccaactagtgttaagcacaccttgactcaccagaagaacttcttcaacaacactctcgaccgcc
 aaggccccaacctccaactcagactacctctgtagttgagtcagccaaggcggtgaaataacctgtcccc
 ctgtagcatccaagggtatcctcgaagagtgtctcccaagaagggtaccaagaagaccactacgaagaagg
 ctcccaaccaatcgcaagaaggcgcatagtgtctactcatcgctcgatgccgccatggaagttaccgctcgtggc
 cactgcaccaactaa

FIG. 19A**SEQ ID NO:38****Protein sequence of Fv30A**

mlfslylptlafqaslalqdtstvtvdtstqklqvidgfgvseayghakqfqnlpgppqkegidllfntttg
 aglsiirnkiqcdasnsitstntdnpdkgavyhfdgdddqgsaqsmgrlccgtpgvscssgdwrhryvemi
 aeylsyykqagipvshvgflnegdgsdfmlstaeqaadvipllhsalqskglgdikmtccdnigwksqrd
 ytaklaelevekylsvitsheyssspnqpmnttlptwmsegaandqafatawyvnggsnegftwavkiaq
 givnadlsayiywegvetnnkgsishvidtdgtkftissilwaihwsrhirpgahrlstsgvvqdtivg
 afenvdgsvvmvltnsqtaaqtvdlgvsgssfstafaftsdacqmvdtkvltlsdgrvkvtvpvhgvvtv
 klttaksskpvtavsaqsaptptsvkhtlthqktssttlstakaptstqttsvvesakavkypvppvas
 kgssksapkkgtkktttkkgshqshkahsathrrcrhgsyrrghotn

FIG. 19B

SEQ ID NO: 39

Nucleotide sequence for Fv43F, a GH43 family enzyme from *Fusarium verticillioides*

atgltggaaaacccctcgctcagcggtctcttgctgcgcgtcgcgctccctcagcgggcggtgaacgctgcttatcccta
accctgggtcccgctcaccggcgatactcgtggttcacgacctacgggttggtcaagaactcccagcggtgggata
cttgctgggtcatactggcgataacggtttcgctcaagaacttcttctcgatcgaaactgcttggaaggatgca
gggtgctggttttccccaaacgggtgcgcccttggaactacgcagtaaccaaggggcgacaagaacctctgggcc
ctgatatctcctaccacaacgggcagtaactatctgtactactccgcctcttcccttcgggtcagcgtaacctc
tgccattttttctcgtctaccagcaagacccgggtgcatccgggtctcgtggaccaaccaaggcgctcgtcgtcgag
tccacaacaacaacagactacaatgccattgacggaaatctcttttgcgactctgatggaaaatgggtggc
tctccttcgggtctctttctgggtccggcatcaagctcatccaactcgcaccccaagacccggcaagcgcaaccgg
ctcaagcatgtactccctcgccaaaacgcgaacgcctccgtcgaaagcgccgctcgaggctccggttcacacc
aaacgcgggaagcaactactacctctgggtggtggttcgacaagtggtgcccaggggcgctgctagcagctacc
gtgtcatgggttggaacgggtcgagcagcattactgggtcccttatggttgacaaggttggttaagcagatgatgtc
tggtggagggaacgggagattatggctagtcacgggatctattcatggacccgggacataatgctggttttccact
gataacgatgoggacggttcttggtctatcattactacgataacgctggcacagcgctggttgggcatcaact
tgctcagatatgacaatggctggcctggttgcttattag

FIG. 20A

SEQ ID NO:40

Protein sequence for Fv43F

mwklllvsglvavaslsqvnaaypnpgpvtgdtrvhdpvtvvktpsggyllahtgdnvslktsdrtawkda
gavfpngapwttqytkgdkniwapdisyhngqyylyysassfgqrtsaiflatsktgasgswtnqgvvve
snndynaidgnlfvdsdgkwwlsfgsfwsgikliqldpktgkrtgssmyslakrdasvegaveapfit
krgstyyllwvsfdkccqgaastyrmvgrsssitgpyvdkagkqmmsgggteimashgsihgpghnavft
dndadvlvhyhynactailqinllrydnqwpvay

FIG. 20B

SEQ ID NO:41

Nucleotide sequence for Xyn3, a GH10 xylanase from *Trichoderma reesei*

atgaaagcaaacgtcatcttgtgctctctggccccctggctgcgcgctctccccaccgaaaccatccacc
 tcgaccccgagctcgccgctctccgcgcacaacctcaccgagcggaacagccgacctctgggacccgccaagc
 ctctcaaagcatcgaccagctcatcaagagaaaaaggcaagctctactttggcaccgccaaccgaccgggc
 ctctccaacgggaaaagaadcgccgccatcatccaggcagacctcgccaggtgacgcgggagaaacagca
 tgaagtggcagtcgctcgagaacaaccaaggccagctgaactggggagacgcccactatctcgtaactt
 tgcccagcaaaacggcaagtcgatacgcggccacactctgatctggcactcgcagctgcctgcgtgggtg
 aacaatatcaacaacgcggatactctgcggcaagtcacccgcaacctgtctctactgtggttggggcgt
 acaagggaagattcggtgcttgggtgagttttgaacaccacatgccccctttcttagtccgctcctctc
 ctcttggaaacttctcacagttatagccgtatacaacattcgacaggaaatttaggatgacaactactgac
 tgacttgtgtgtgtgatggcgataggacgtgggtcaatgaaattctcaacgaggatggaacgctgcgctct
 tcagctcttttccaggctcctcgccgaggagtttgtctcgattgcctttcgtgctgctcgagatgctgacc
 cttctgcctgctctttacatecaacgactacaatctcgaccgcgccaactatggcaagggtcaacgggttgaa
 gacttaagctctcaagtggtatctctcaaggagttccattgacgggtattggtgagccacgacccctaaat
 gtcccccattagagctctctttctagagccaaggcttgaagccattcagggactgacacgagagccttctc
 tacaggaagccagtcctcatctcagcggcggcggaggctctggtaacgctgggtgcgctccagcagctggca
 acggtaaccgctcaccgagctggccattaccgagctggacattcagggggcaccgacgaacggattacaccc
 aagttgttcaagcatgcctgagcgtctccaagtgcgtcggcaccacgctgtggggcattcagtgacaaggt
 aagttgcttccccctgtctgtgcttatcaactgtaagcagcaacaactgatgctgtctgtctttacctagg
 actcgtggcgtgccagcaccaacctctctctgtttgacgcaaaacttcaaccccaagccggcatataacag
 cattgttggcatcttacaatag

FIG. 21A

SEQ ID NO:42

Protein sequence for Xyn3

mkanvilcllaplvaa**l**ptetihldpelaalranltertadlwdrqasq**sidqli**krkgklyfgtatdrg
 llgreknaa**ii**qadlgqvtpensmkwqslennqgqlnwgdadylvnfaqqngksirghtliwhsqlpawv
 nninnadt**lr**qvirthvstvvgrykgkirawdvvneifnedgtl**rssv**fsrllgeefvsiafraardadp
 sarlyindynldranygvngl**ky**tvskwisqgvpidgigsqshlsgggsgtlgalqqlatvpvtelai
 teldi**gg**aptt**dy**tqv**v**qac**ls**vskcvgitvwgisdkdswrastn**pl**lfdanfnpkpaynsivgilq

FIG. 21B

SEQ ID NO:43

Protein sequence of Xyn2, a GH11 family xylanase from *Trichoderma reesei*

mvsftsllaasppsrascrpaaevesvavekrqtiqqgtgynngyfysywndghggvtytngpgggqfsn
 wsnsngnfvggkgwqpgtknkvinfsgsynpnngnsylsvygwsrnplieyyivenfgtynpstgatkge
 tsdgsvydiyrtqrvnqpsiigtatfyqywsvrrnhrssgsvntanhfnawaqqgltlgtmdyqivaveg
 yfssgsasitvs

FIG. 22

SEQ ID NO:44

Protein sequence of Bxl1, a GH3 family β -xylosidase from *Trichoderma reesei*

mvnnaaallaalsallptalagnnqtyanysaqqgpdlpetlatltlsfpdcehgplknnlvcddssagv
 eraqalislftleelilntqnsqpgvprlgplpnyqvwnealhgl dranfatkggqfewatsfpmpiltta
 alnrtlihqadiistqarafnsgrgldvyapnvngfrsplwgrgqetpgedafflssaytyeyitgi
 qggvdpehlkvaatvkhfagydlennwnnqsrllgfdaiitqgdlseyytpqflaaaryaksrslmcaynsv
 ngvp scansfflqtllreswgfpewgyvssdcdavynvfnphdyasnqssaaasslragtdidogqtypw
 hlnefvagevsrgeiersvtrlyanlvrlgyfdkknqyrslgwkdvvktdawnlsyeaavegivilknd
 gtlplsckvrsialigpwanattqmqqnyygpapylispleaakkagyhvnfelgteiagnsttgfakai
 aaaksdaiiylggidntiegegadrtdiawpgnqldlikqlsevgkplvvlqmggggvdssslksnkky
 nslvwggypgqsggvalfdilsgkrapagrlvttqypaeyvhqfpqndmnlrpdgksnpgqtiwytgkp
 vyefgsglfyttfketlashpkslkfntssilsaphpgytyseqipvftfeaniknsgktespytamlfv
 rtsnagpapypnkwlvgfdrladikpgghssklisipipvsalarvdshgnrivypgkyelalntdesvkle
 felvgeevtienwpleeqqikdatpda

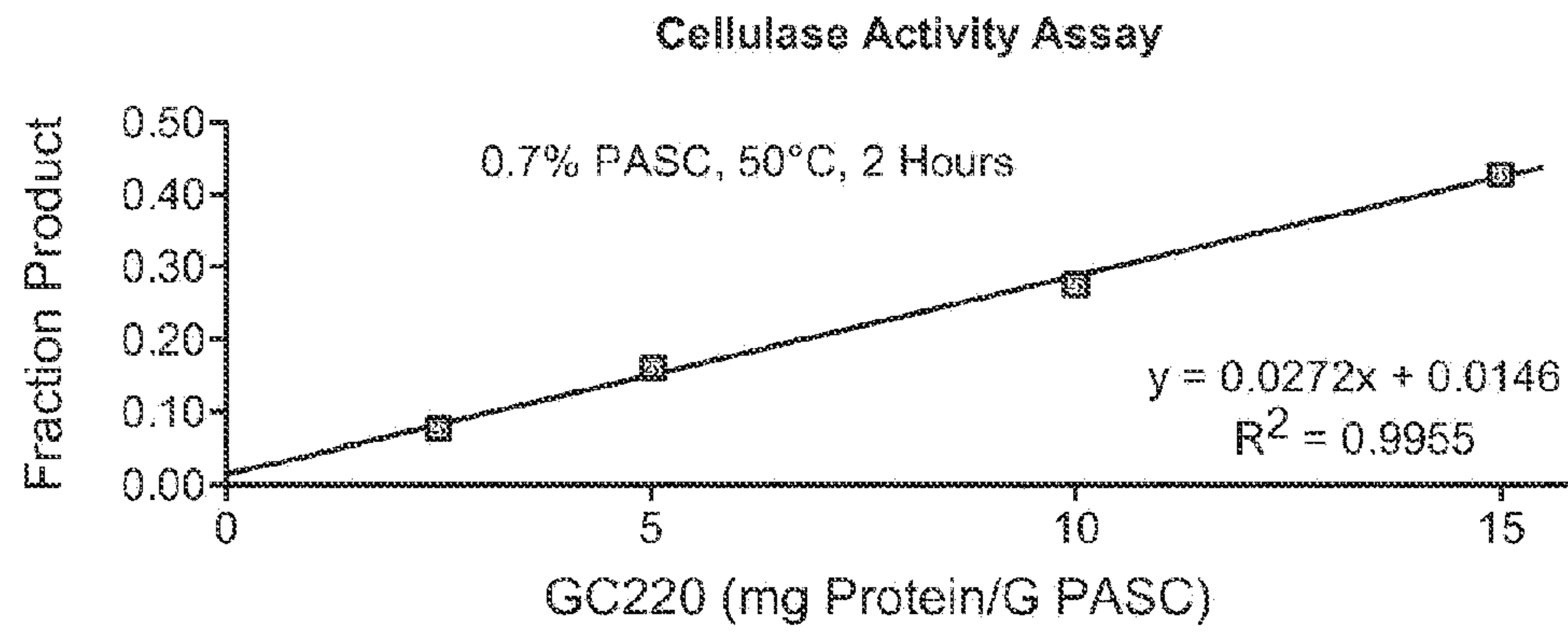
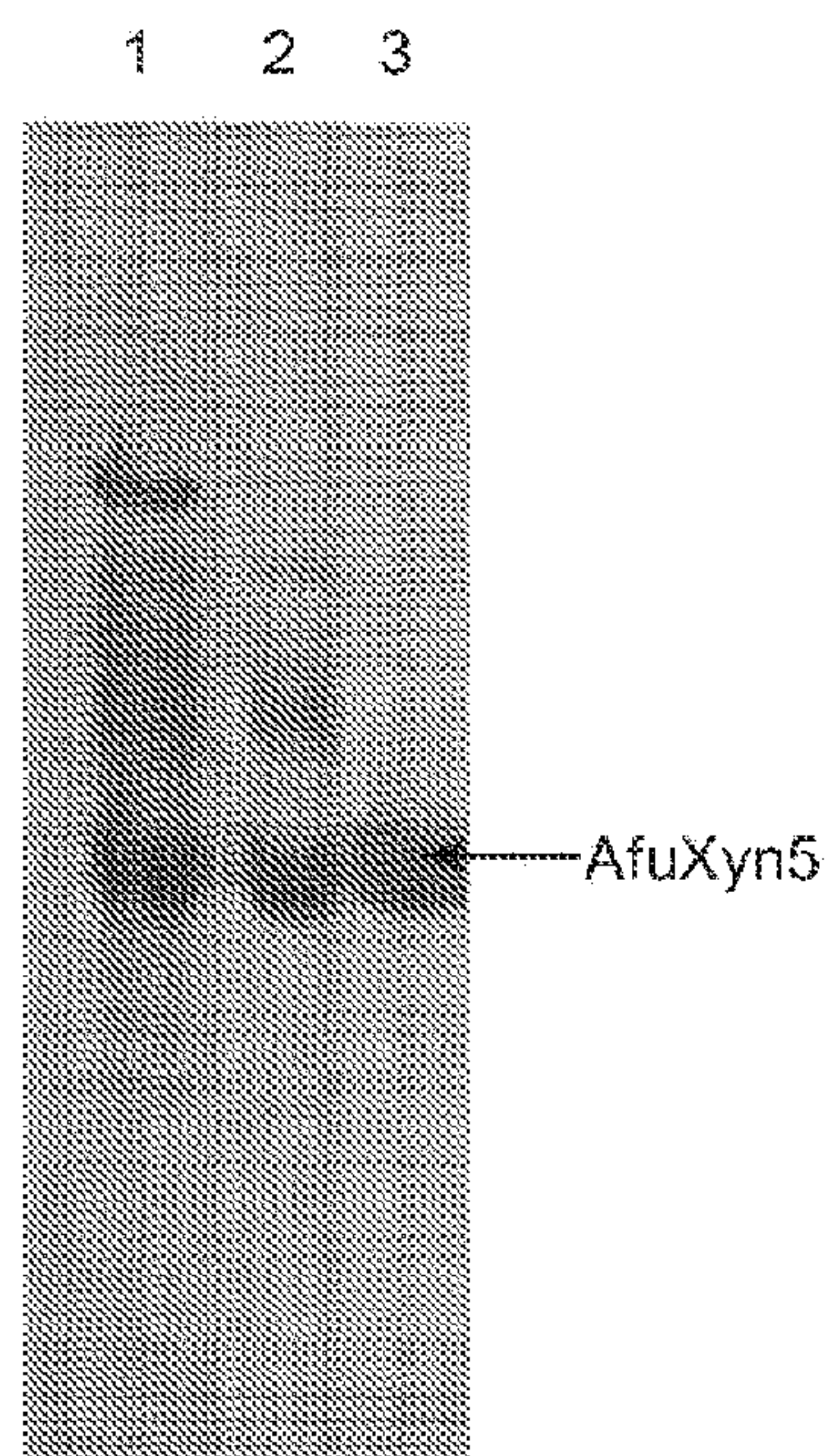
FIG. 23

SEQ ID NO:45

Protein sequence of Bgl1, a GH3 family β -glucosidase from *Trichoderma reesei*

mryrtaaalalatqpfaradshstsgasaeavvppagtpwgtaydkakaalaklnlqdkvgivsgvgwng
 gpcvgnvspaskisypslclqdgplgvystgstafpvgvqaastwdvnlirergqfigeevkasgihvi
 lgpvagplgktpqgggrnwegfgvdpyltgiamgqtingiqsvgvqatakhyilneqelnretissnpddr
 tlhelytwpfadavqanvasvmcsynkvnttwacedqytlqtvtkdqlgfpgyvmtdnaqhttvqsans
 glmsmpgtdfngnnrlwgpaltnavnsnqvptsrvdmdvtrilaawyltgqdgagypsfnisrnnvqgnh
 ktnvraifardgivilkndanilplkkpasiavvgasaaiignharnspscndkgcddgalgmwgwsgavny
 pyfvapydaintrassqgtqvtlsntdntssgasaargkdvaivfitadsgegyitvegnagdrnldpw
 hngnalvqavagansnvivvvhsvgaiileqilalpqvkvavwaglpqsgesgnalvdvlwgdvspsgklv
 ytiakspndyntrivsggsdsfseglfidykhfddanitpryefgyglsytkfnysrlsvlstaksgpat
 gavvpqgpsdlfqnvatvtvdiansgqvtgaevaqlitypssaprtppkqlrgfaklnltpgqsgtatf
 nirrrdlsywdtasqkwvpsgsfgisvgassrdirltstlsva

FIG. 24

**FIG. 25****FIG. 27**

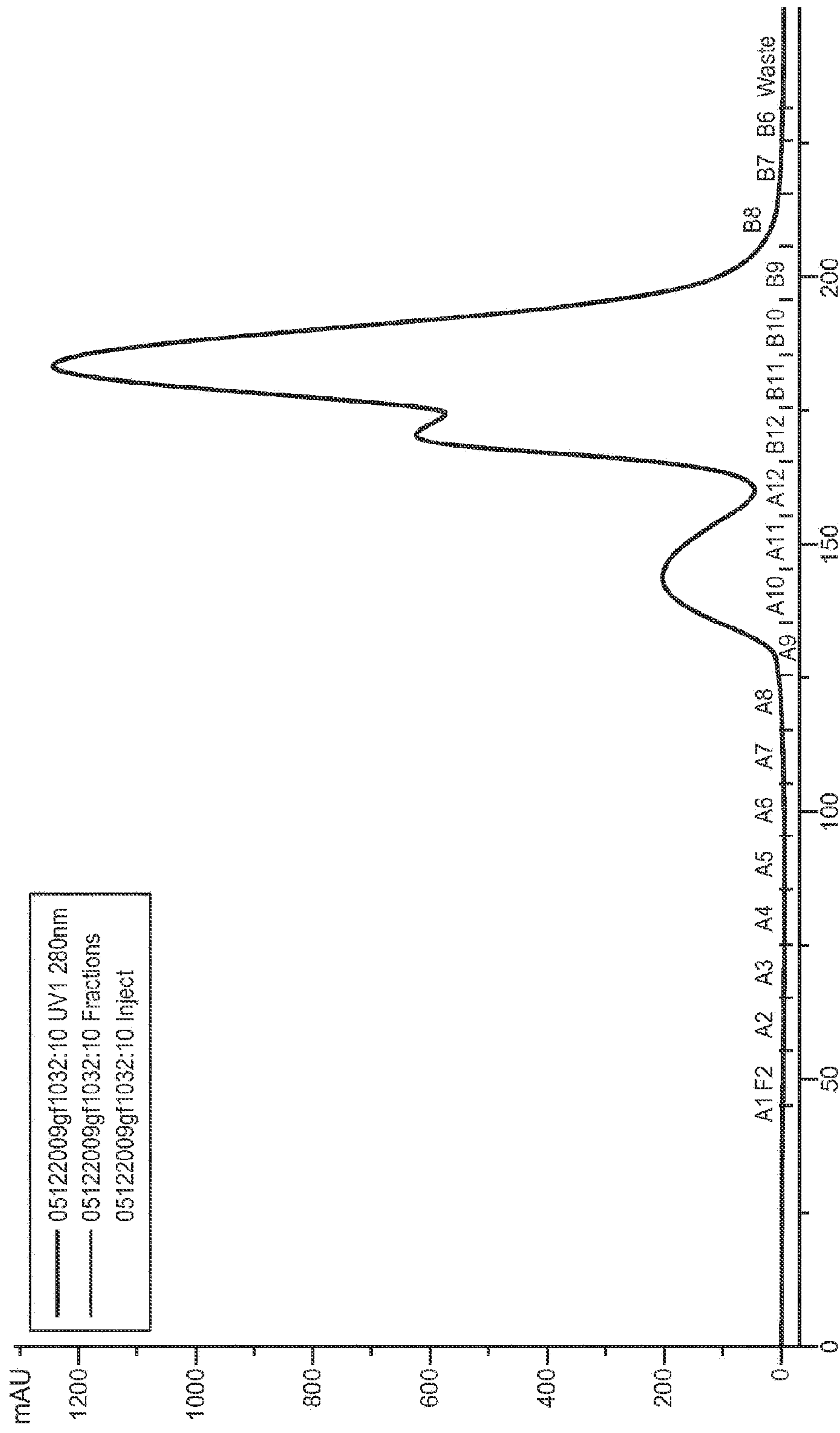


FIG. 26

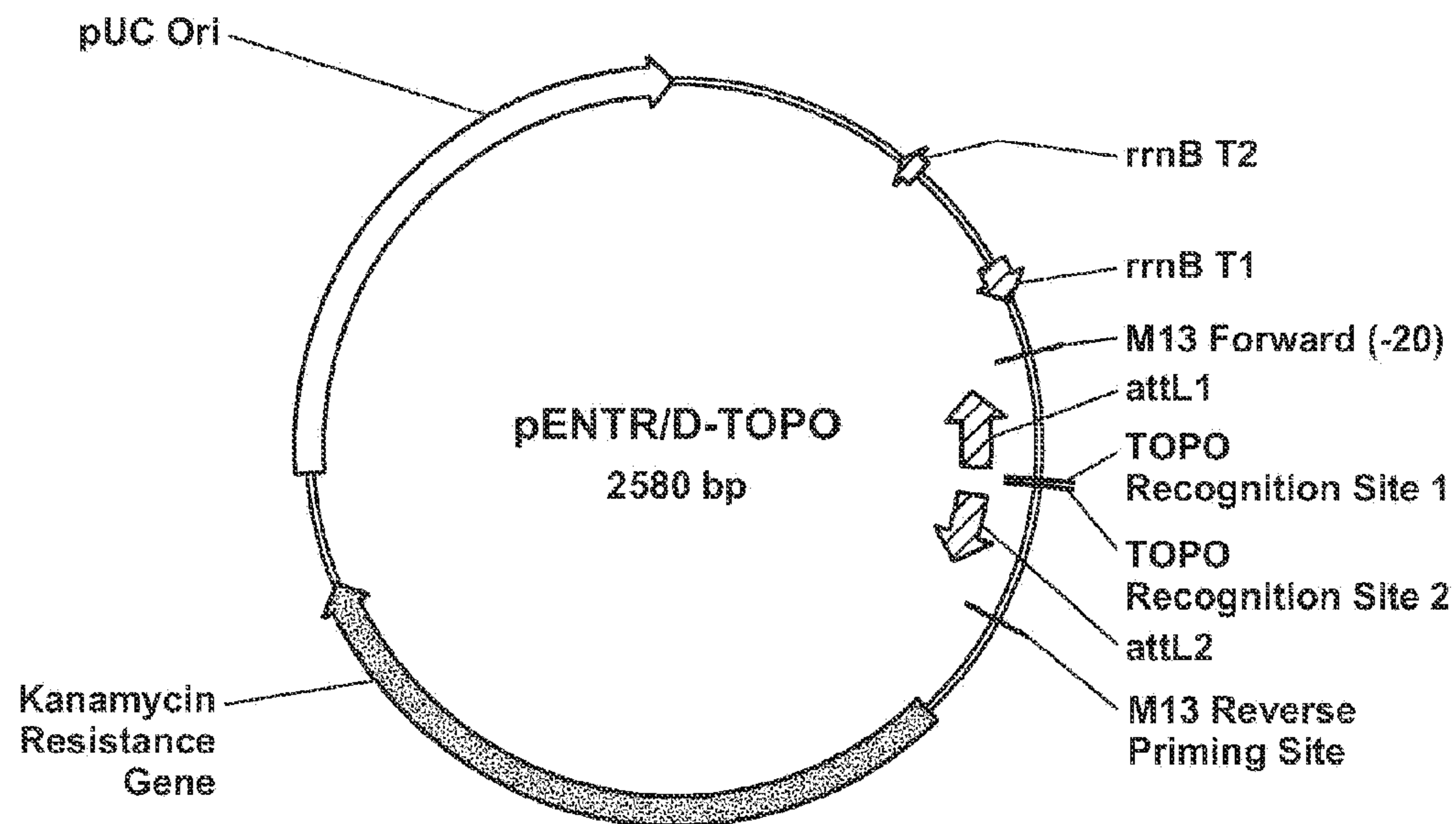
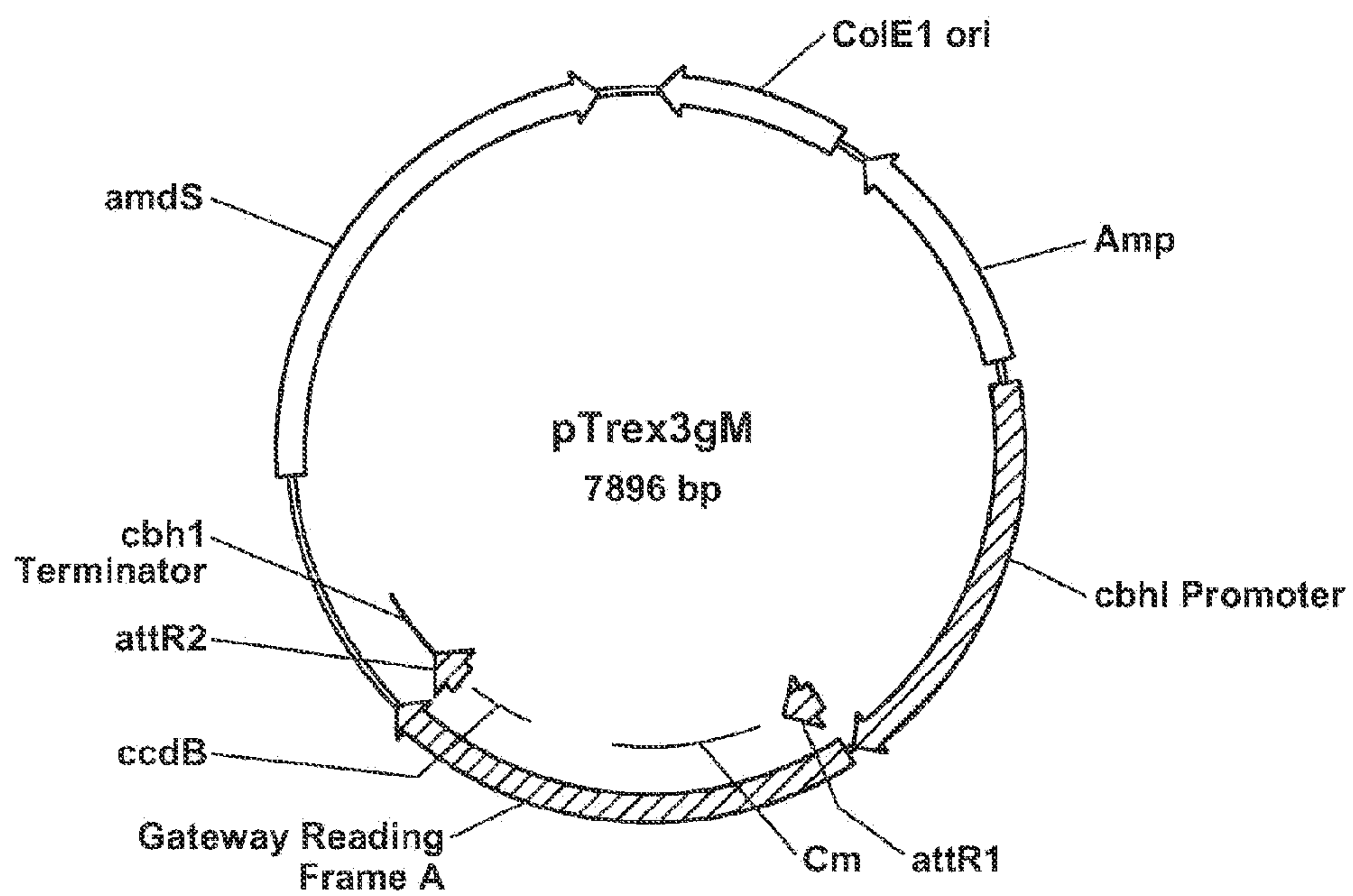
**FIG. 28****FIG. 29**

FIG. 30A

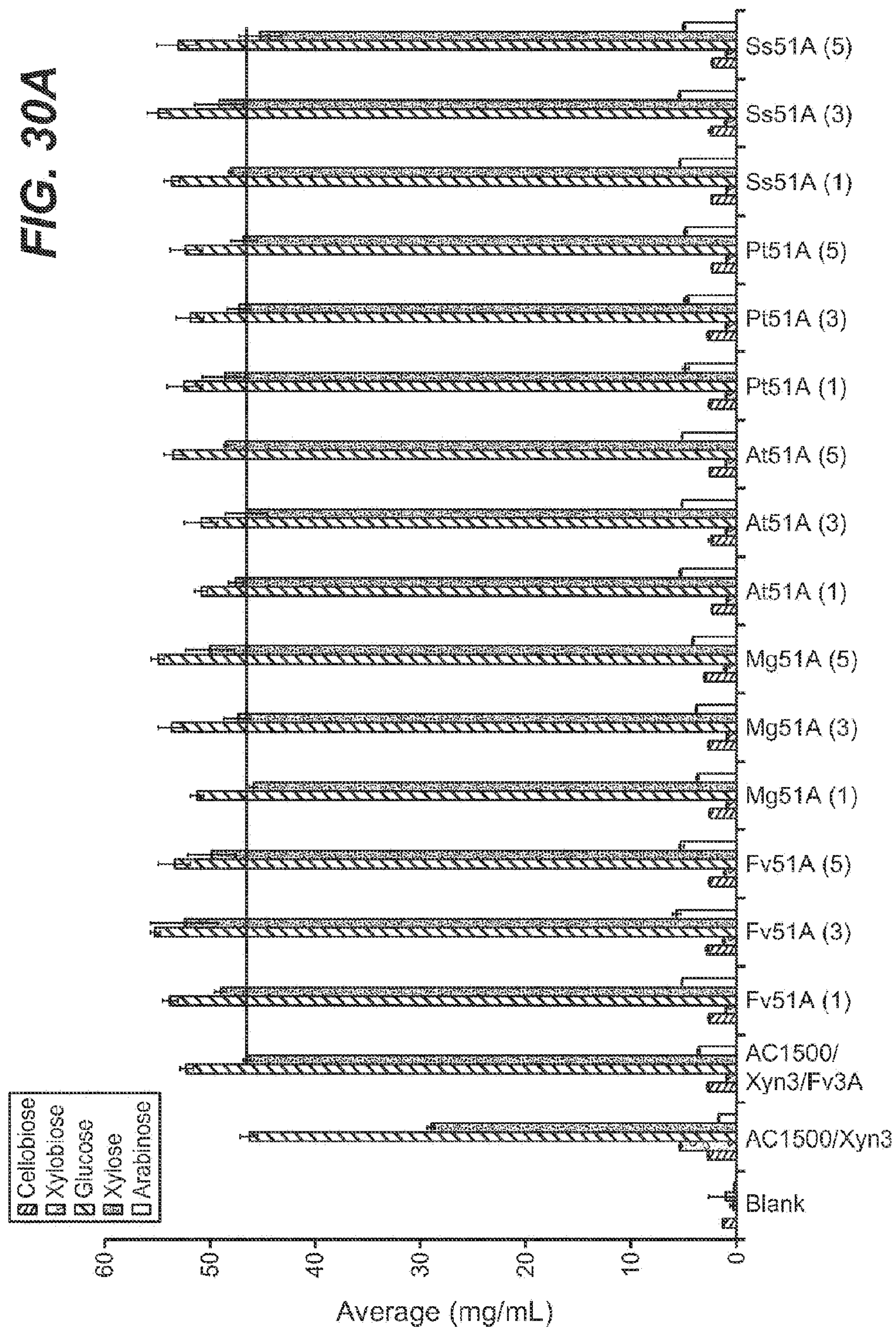


FIG. 30B

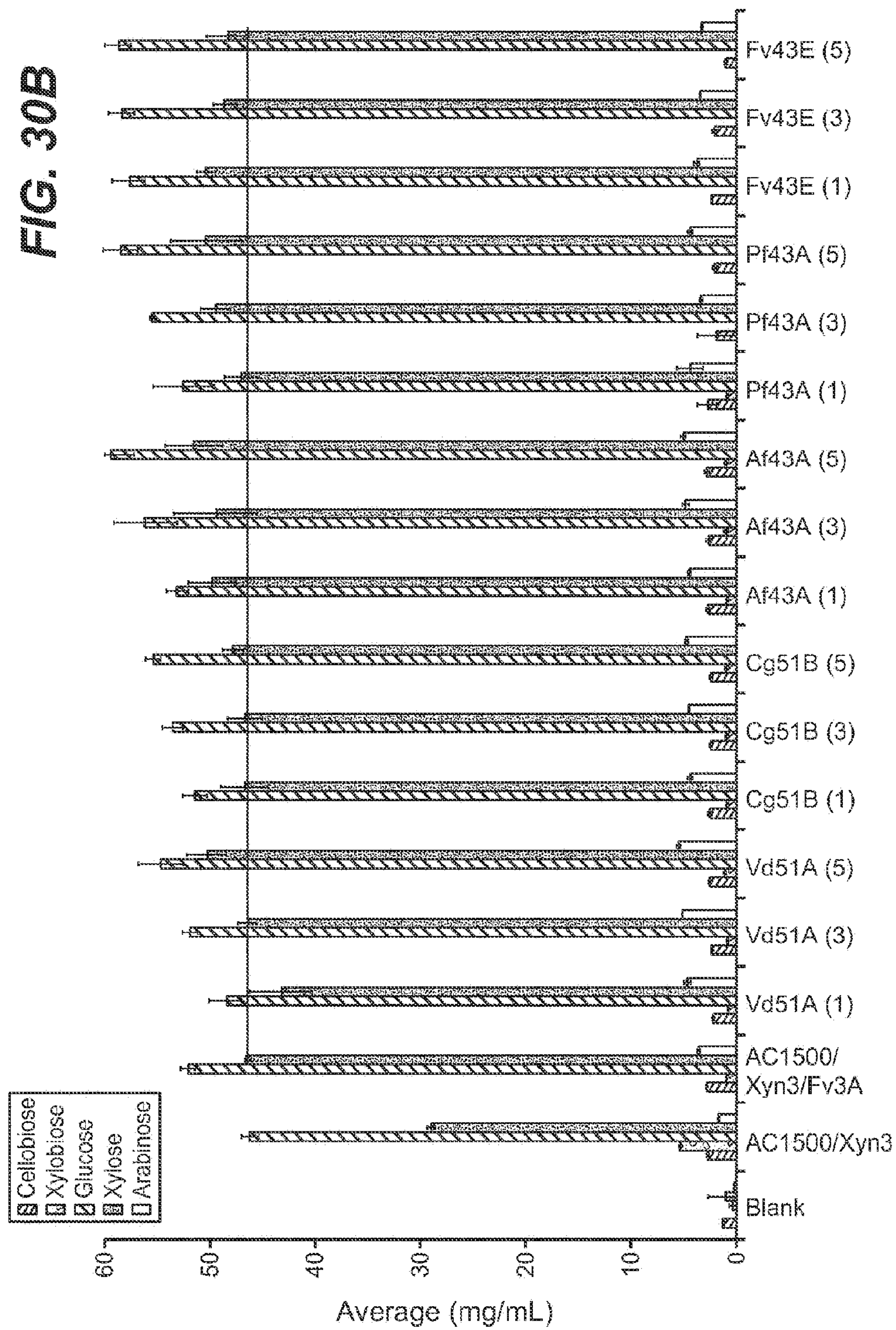


FIG. 31

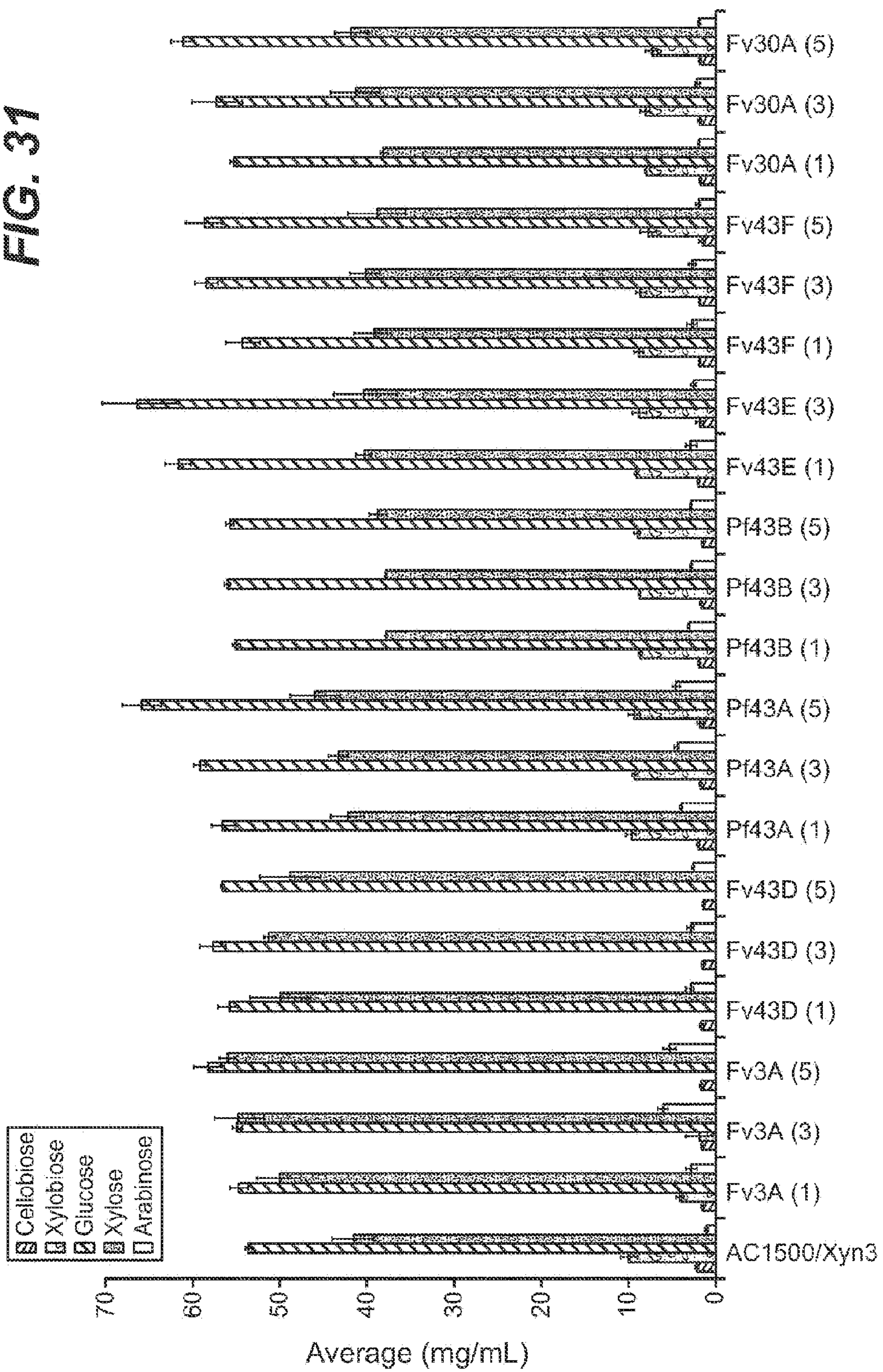


FIG. 32

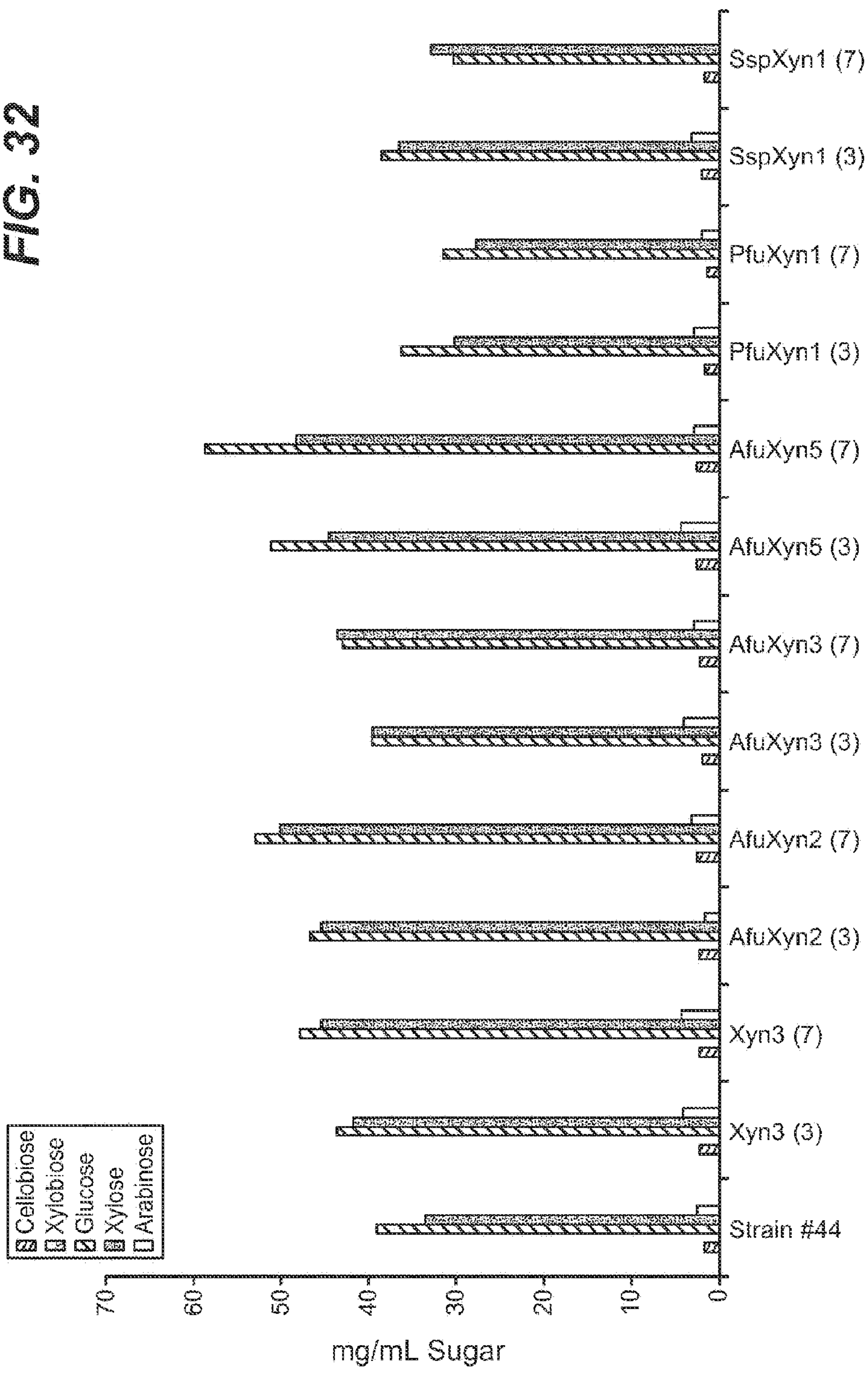


FIG. 33

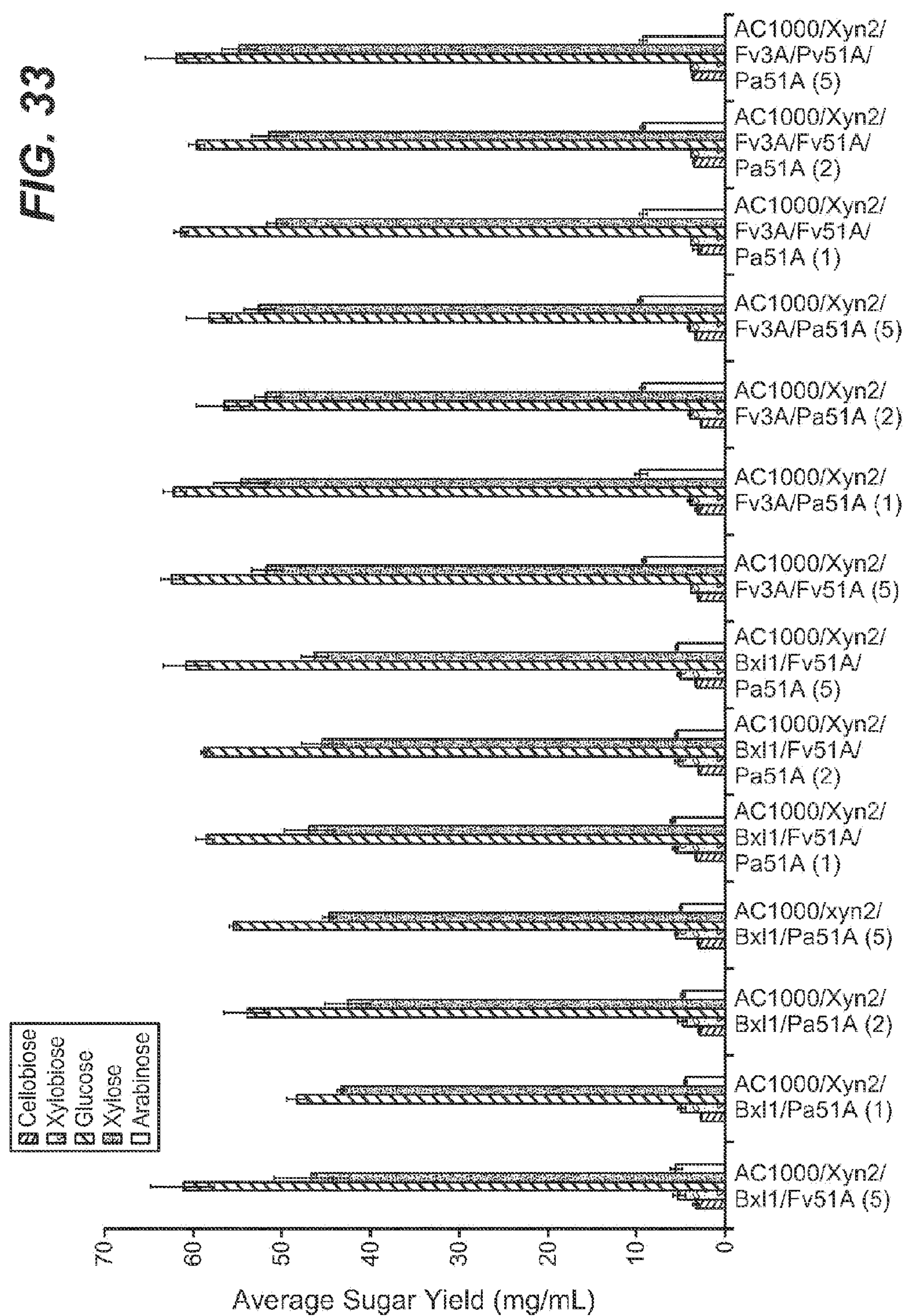
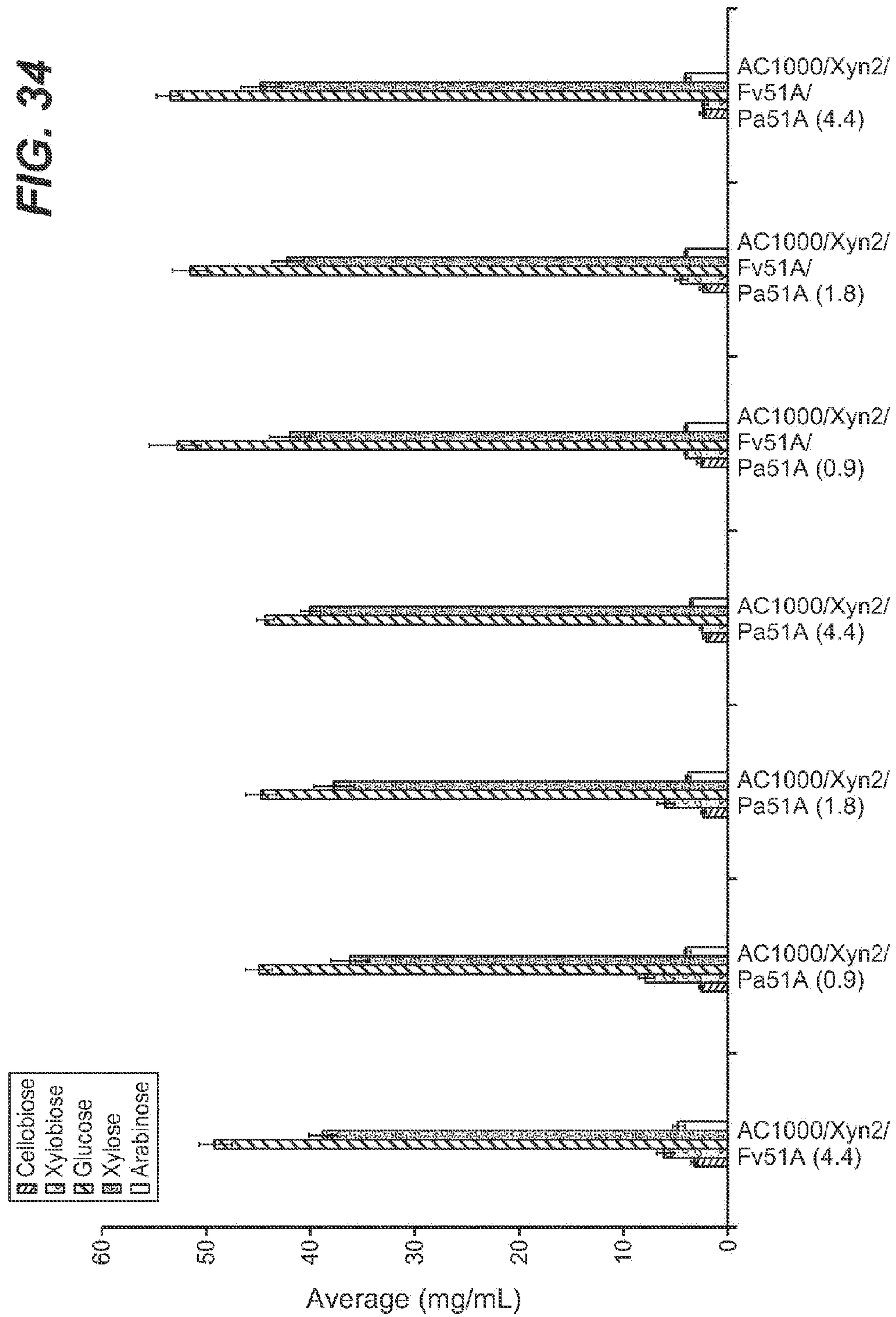


FIG. 34



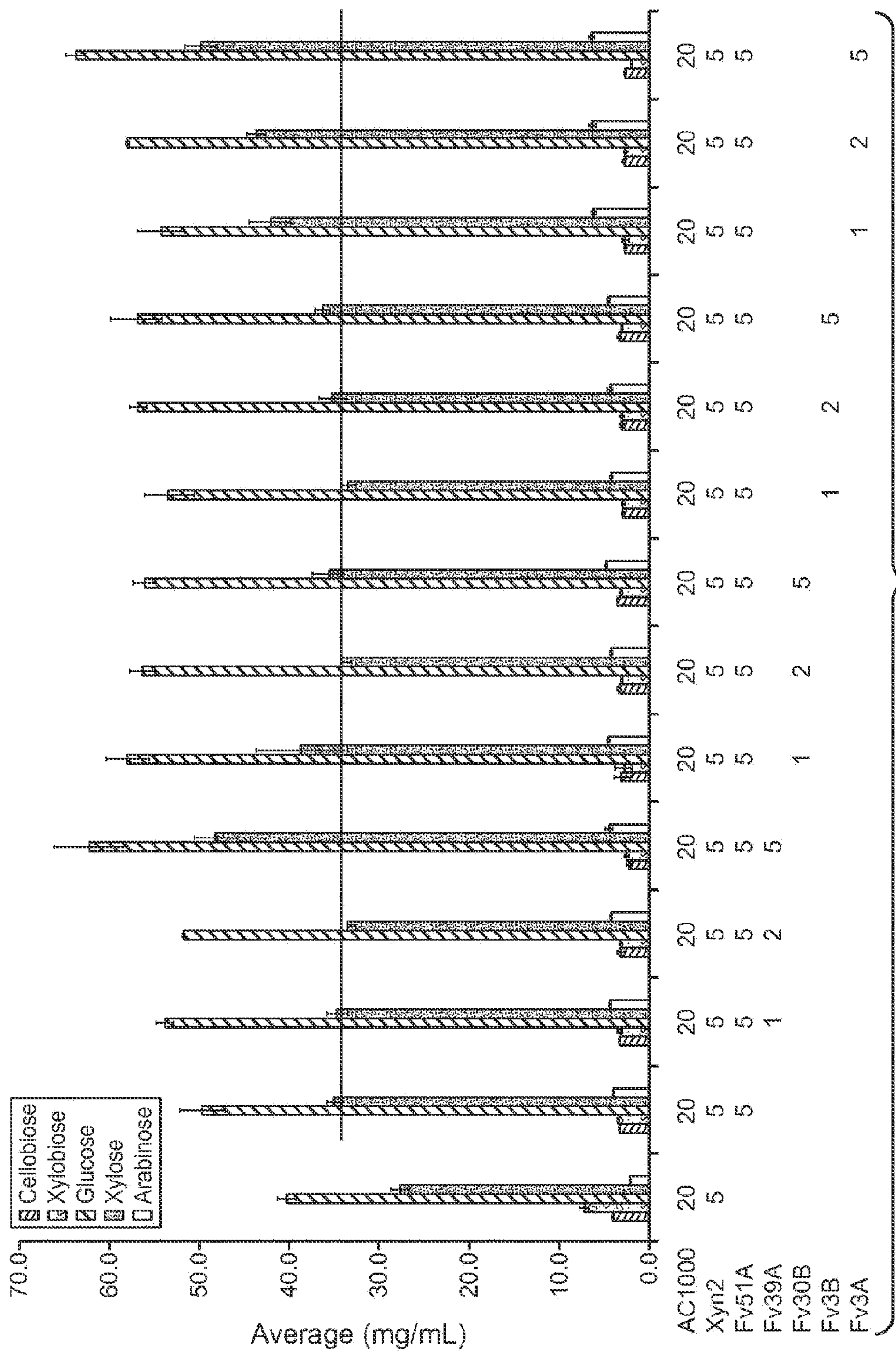


FIG. 35A

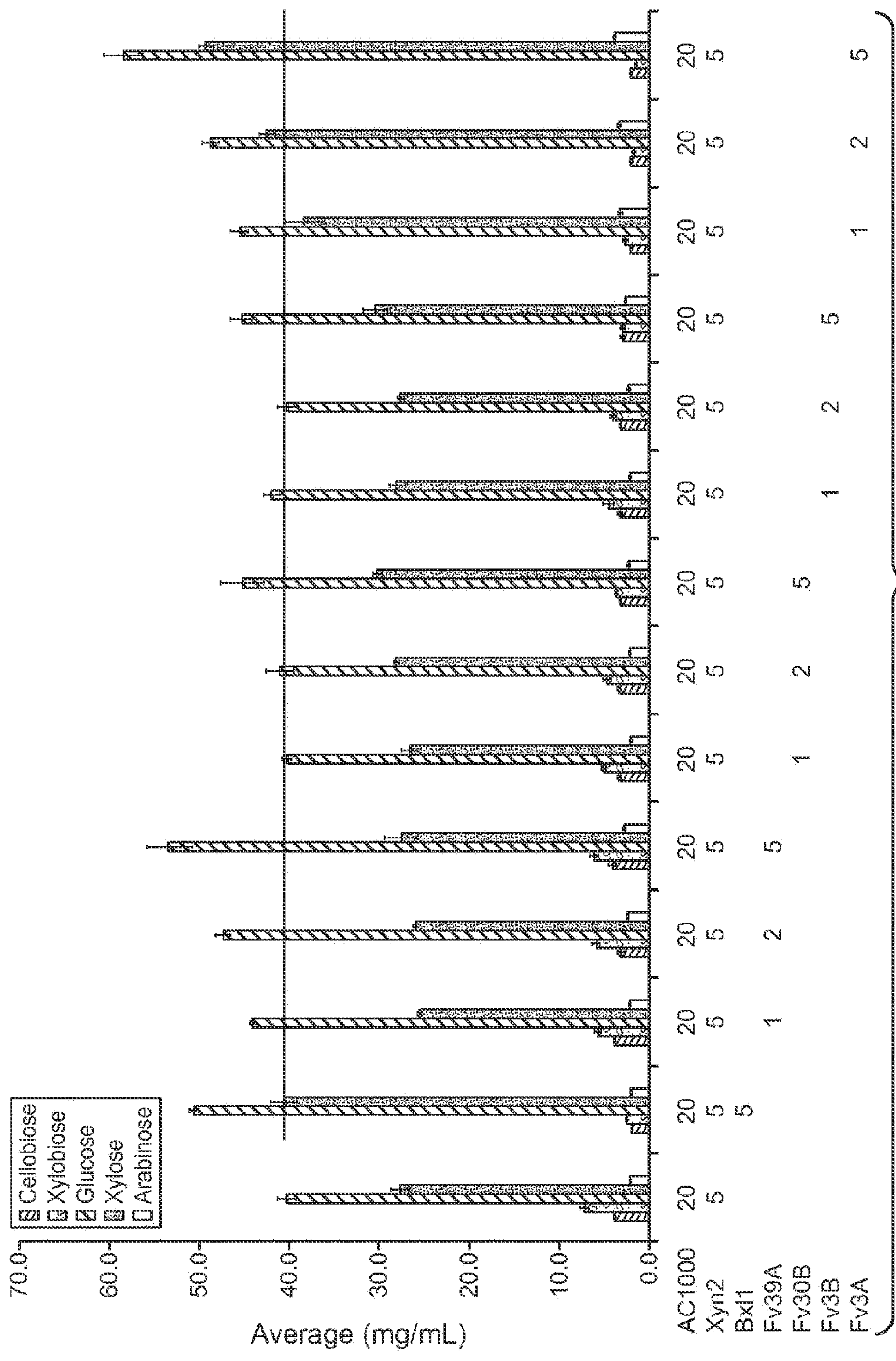


FIG. 35B

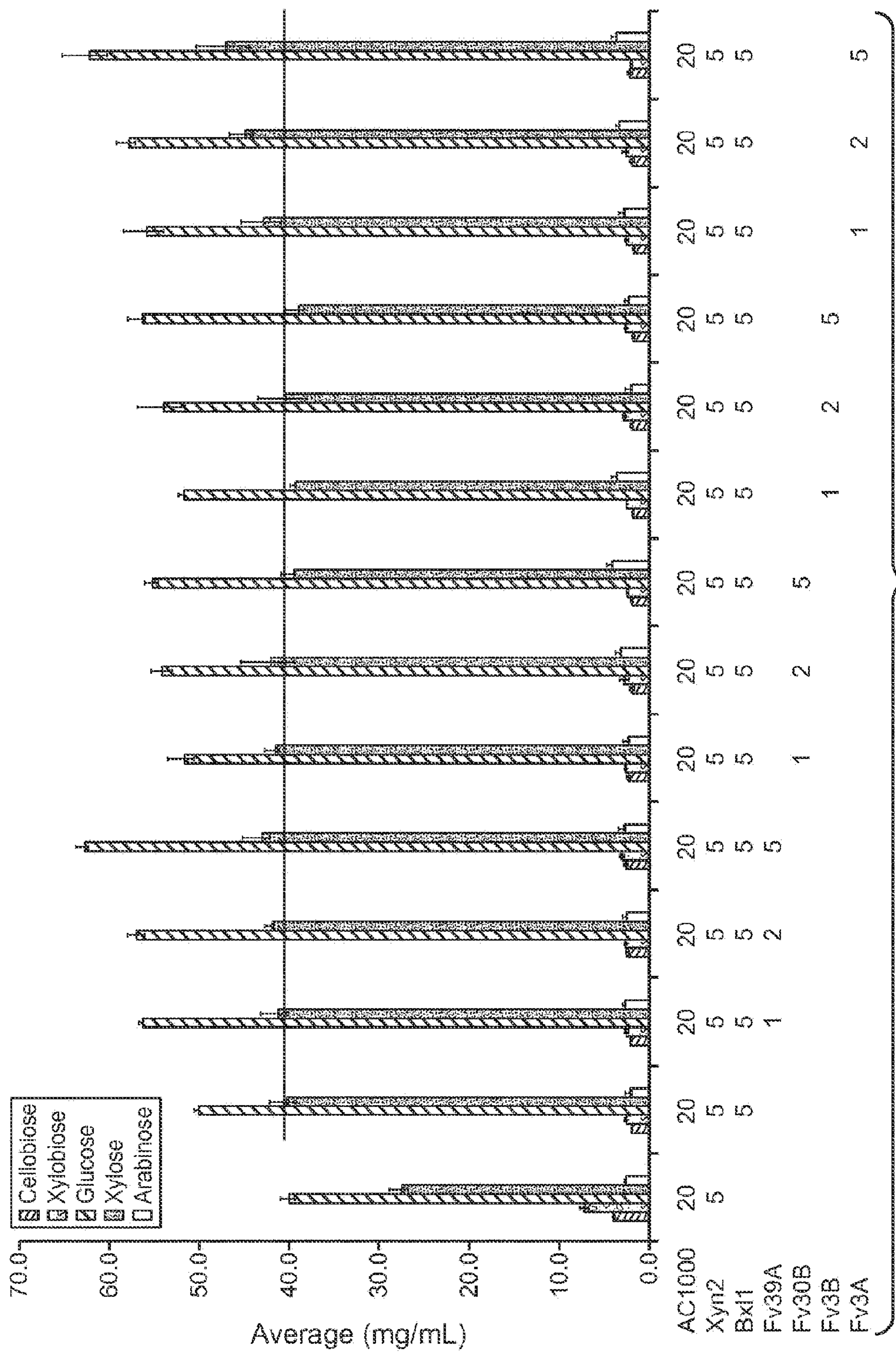


FIG. 35C

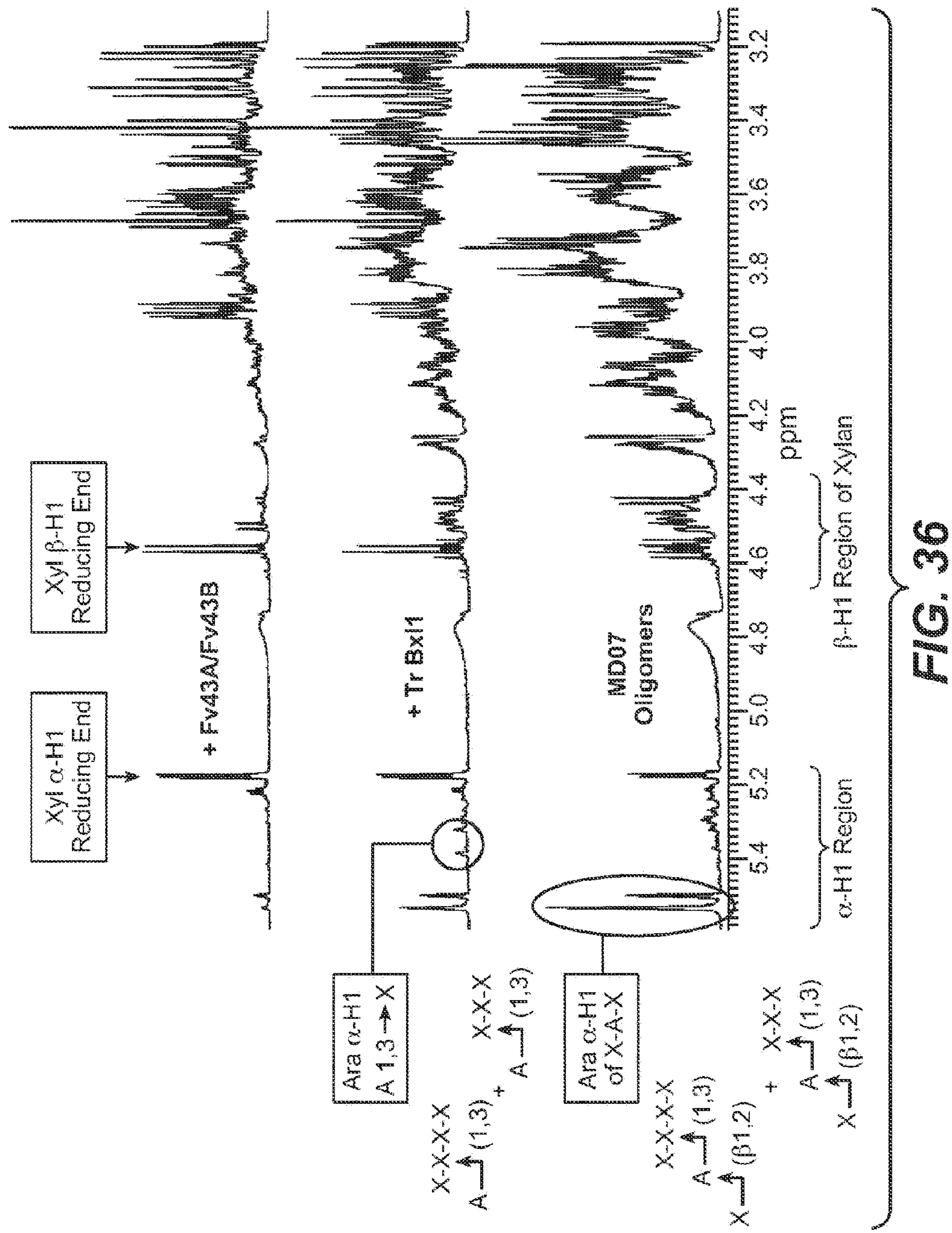
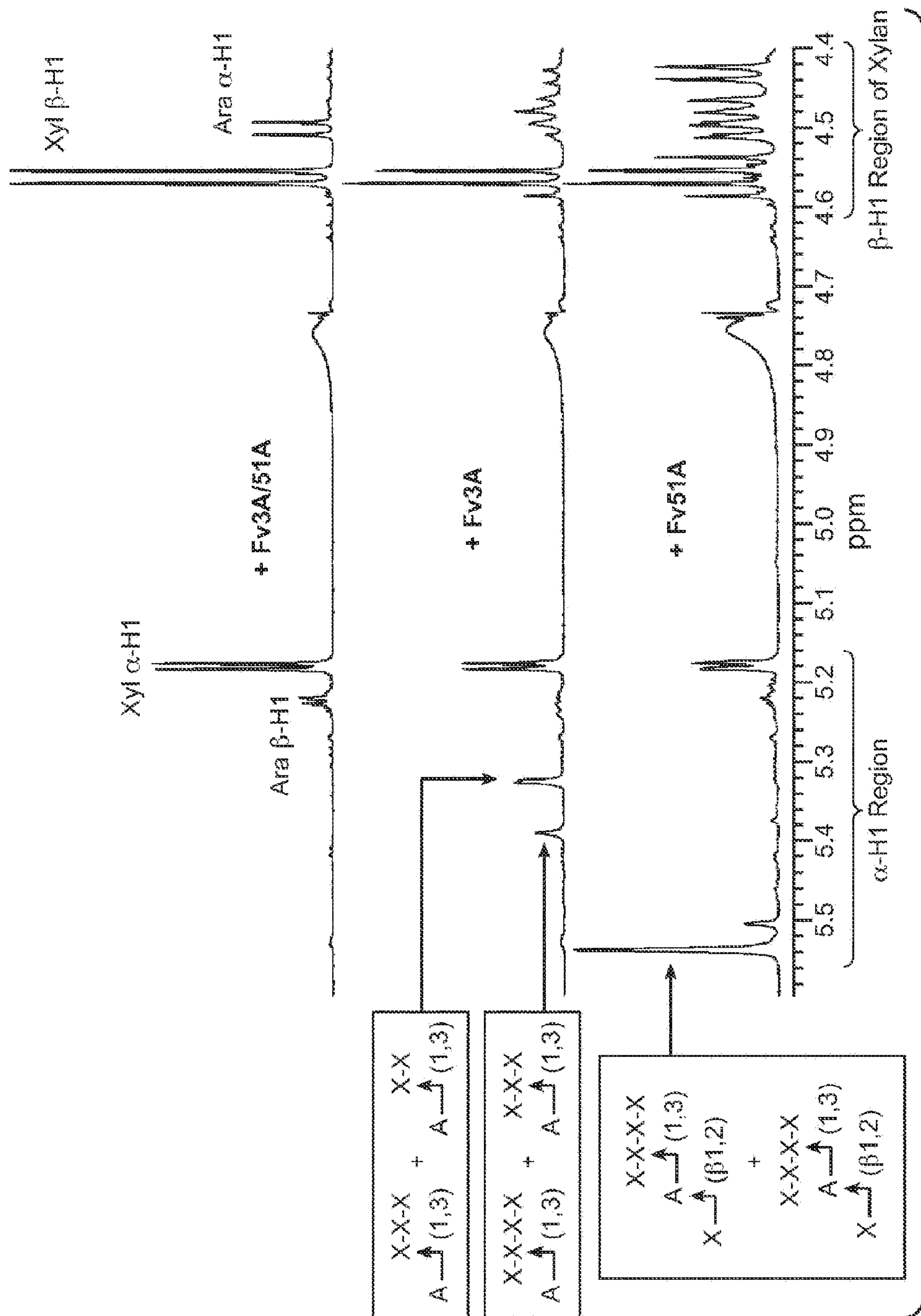


FIG. 36

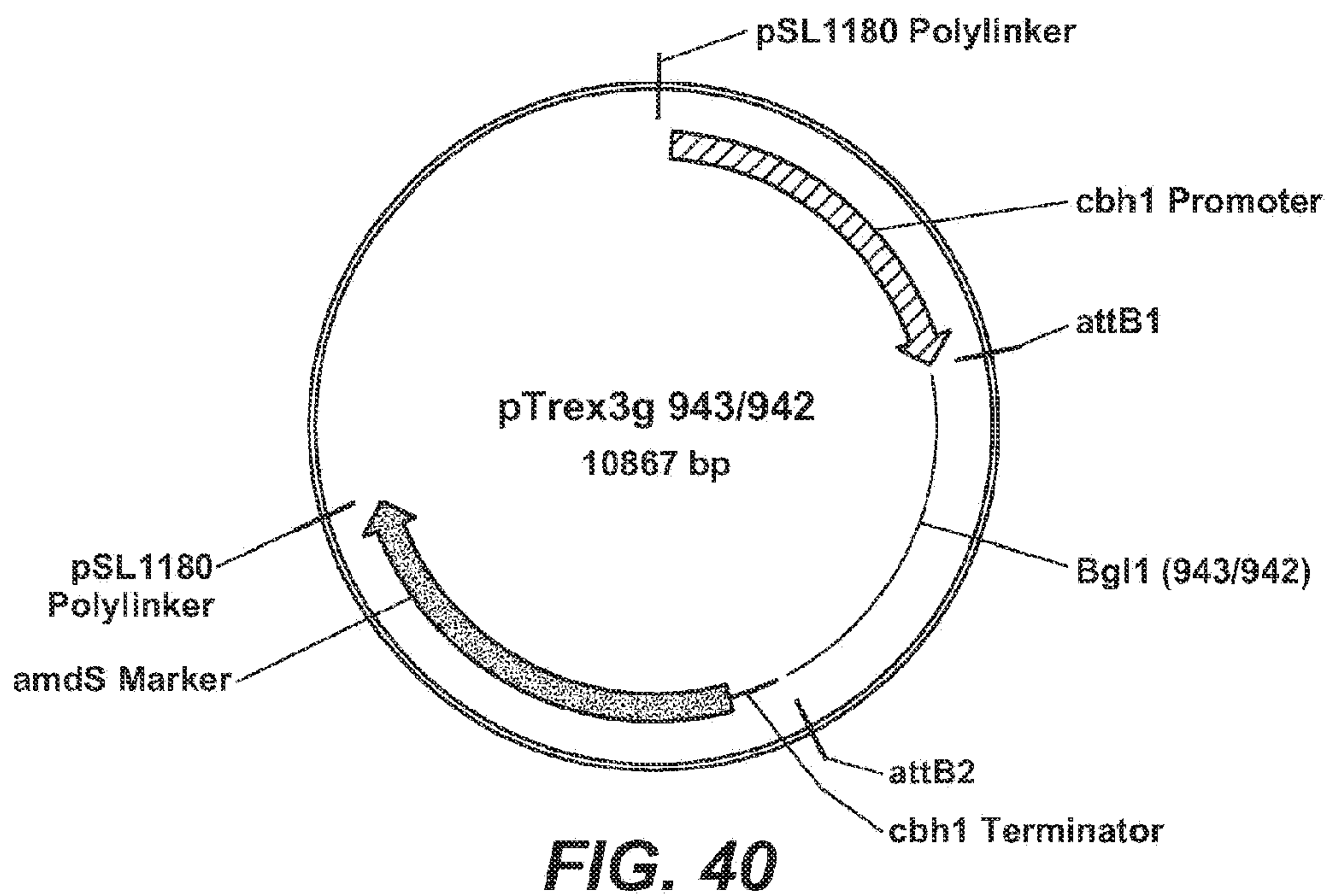
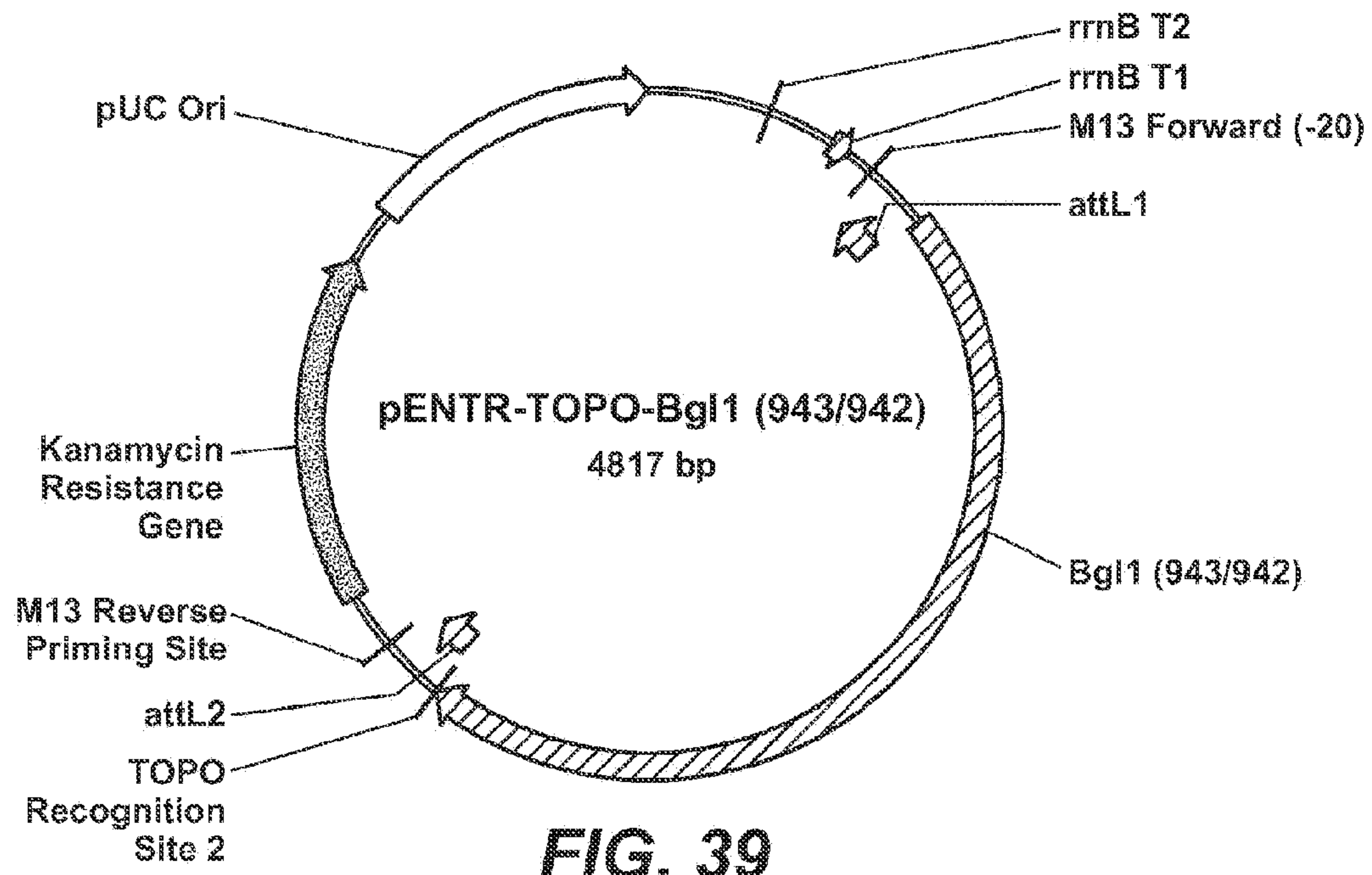


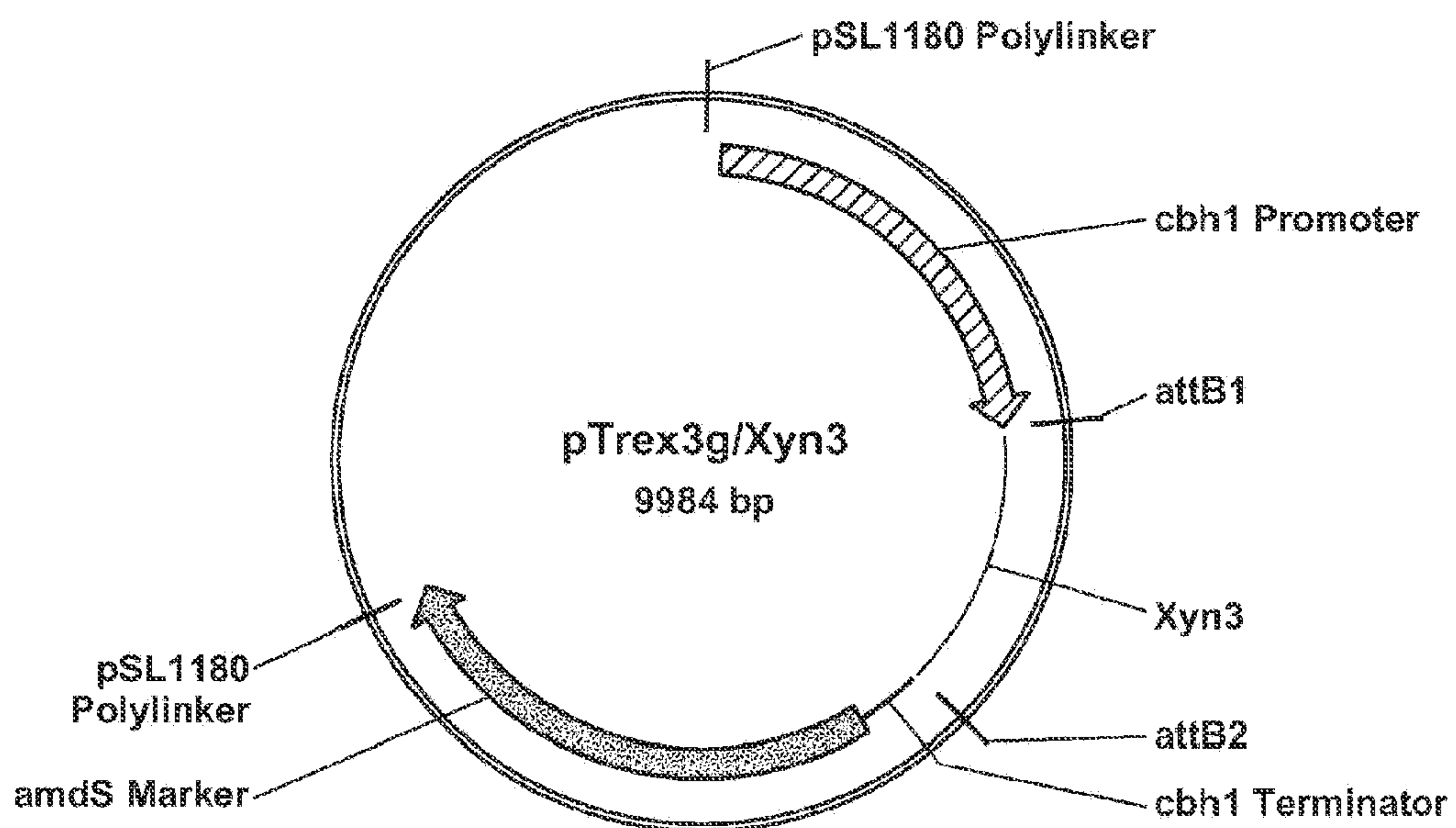
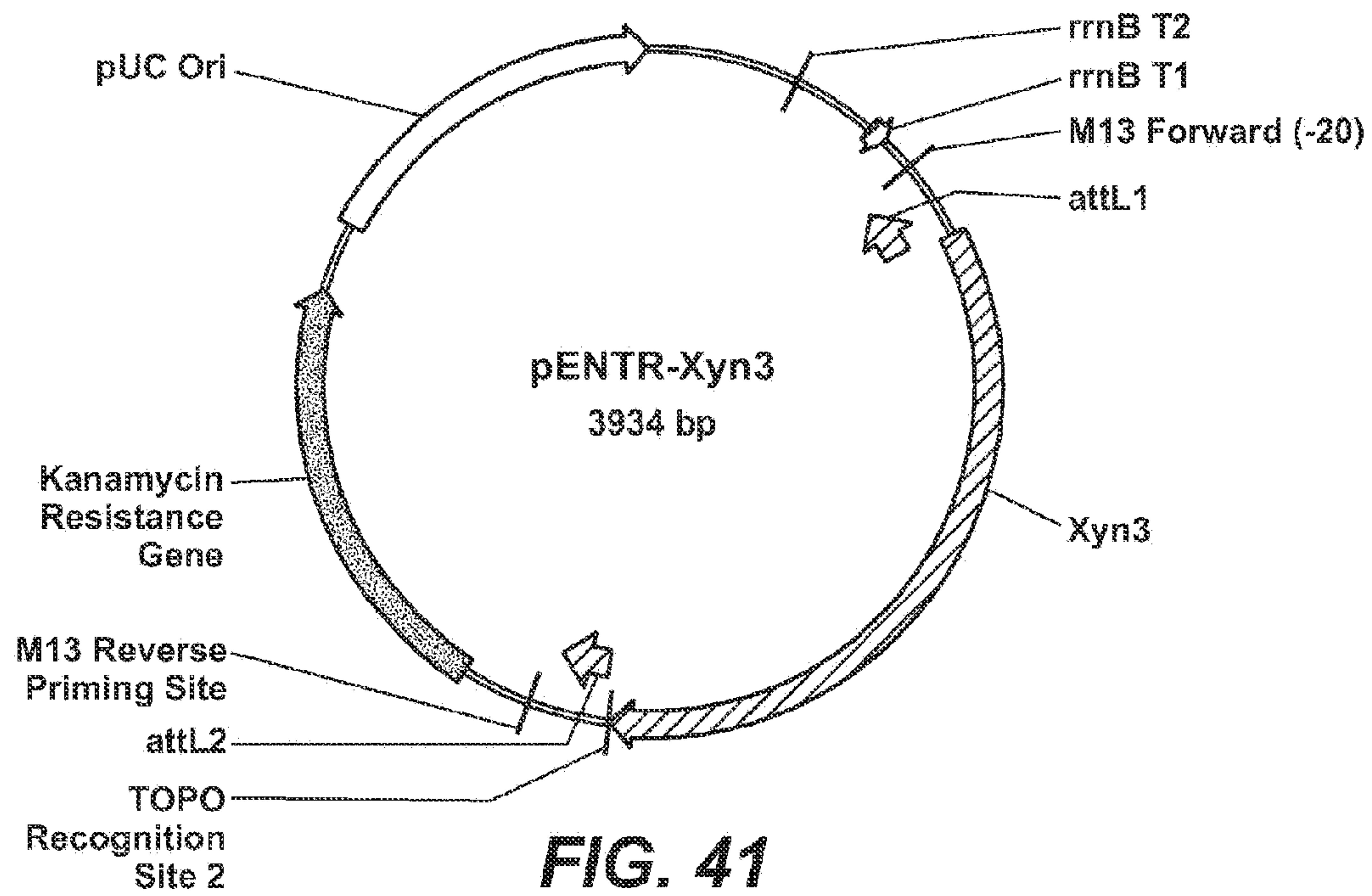
```

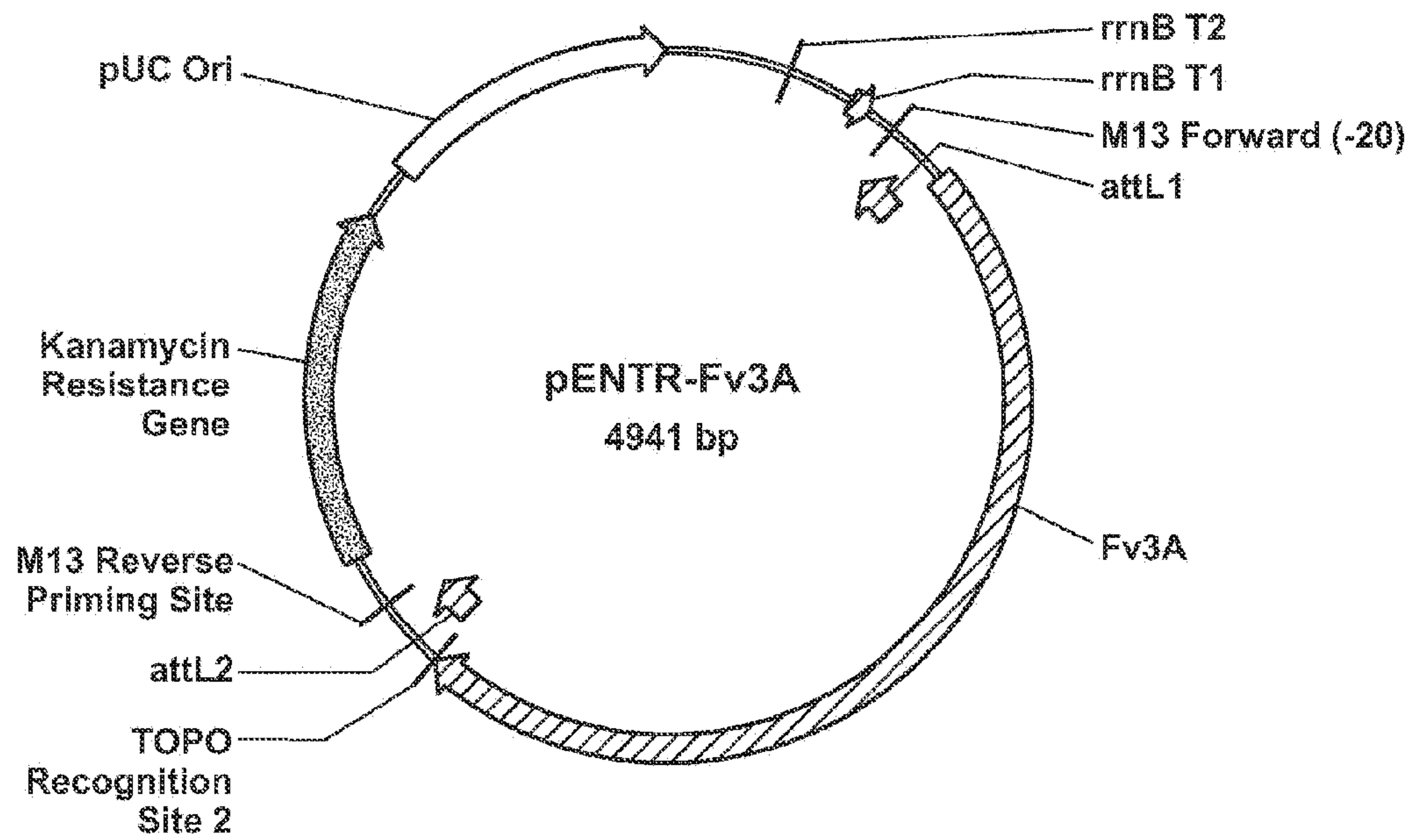
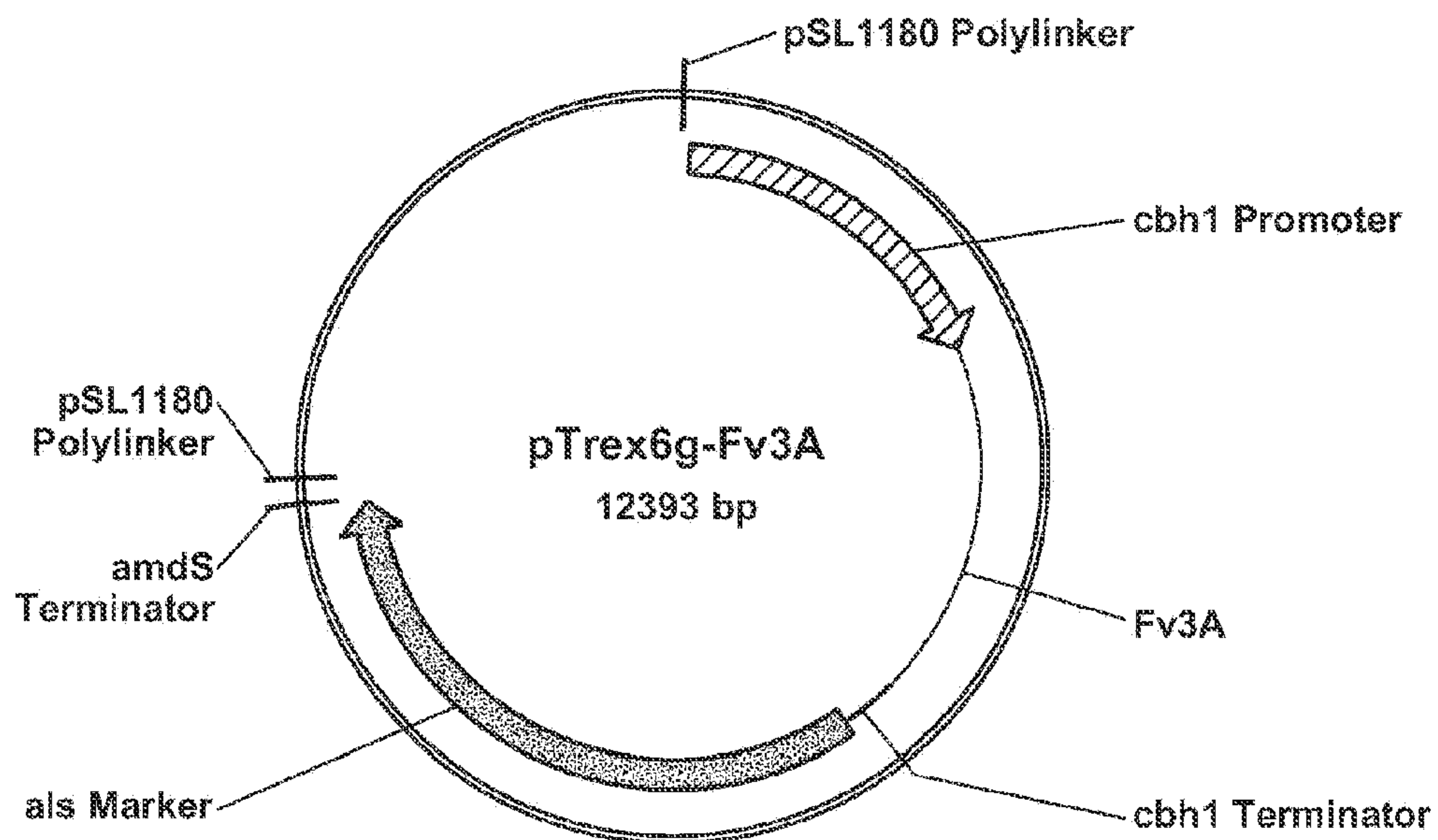
gi|reesei|Bx11 -MVNNAALLAALSALIPTALAQNNQTYANYSAQGPDLYPETLATLTLSFPDCEHGPIKN
gi|fvert|Fv3A MLLNLQVAASALSLSILGGLAEAAATPYT-----LPDCTKGPLSK
      :;* . :*** * ,** : .*: :*** :***.:
gi|reesei| NLVCDSSAGYVERAQALISLFTLEELILNTQNSGPGVPERLGLPNYQVWNEALEGLDRAN-
gi|fvert| NGICDTSLSPAKRAAALVAALTPEEKVGNLVSNATGAPRIGLPRYNWNEALHGLACSPG
      * :***: * , .:*** **: : * ** : * ....*:**:**:*: : ***** :
gi|reesei| --FATKGGQFEWATSFPMPILTAAALNRTLHQIADIISTQARAFSNSGRYGLDVIYAPNV
gi|fvert| GRFADTP-PYDAATSEFMPMLLMAAAFDDDLIHDIGNVVGTEARAFNGGWRGVDFWTPNV
      ** . : : *****:* :***: : ***;* ,: : ,*:***:* ,* * :* ,:***
gi|reesei| NGFRSPLWGRGQETPGEDAFFLSSAYTYEYITGIQGGVDPEHLKVAATVKHFAGYDLENW
gi|fvert| NPFKDPWGRGSETPGEDALHVS-RYARYIVRGLEG--DKEQRRIVATCKEYAGNDFEDW
      * * ,* **** ,*****: ,* * : : * : * * * : ,* ** * : * : * :
gi|reesei| NNQSRLGFDALITQQDLSEYYTPQFLAAARYAKSRSLMCAYNSVNGVPSCANSFFLQTL
gi|fvert| GGFTRHDFDAKITPQDLAEYYVRPEQECTRDAGVGSIMCAYNAVNGIPACANSYLOETIL
      . . :* .*** ** ***:*** . * .: * * * :*****:***:***:***: :*:
gi|reesei| RESWGFF-EWGYVSSDCDAVYNVENPHDYASNQSSAAASSLRAGTDIDCGQTPWHLNES
gi|fvert| RGHWNWTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMDSSCEYTTTSDVSDS
      * * . . : : :***: * : : : * . . . . . * : : * * . * * . . : :
gi|reesei| FVAGEVSRGEIERSVTRLYANLVRLGYFD-KKNQYRSLGWKDVVKTDAWNISYEAAVEGI
gi|fvert| YKQGLLTEKLMDRSLKRLFEGLVHTGFFDGAQAQWNSLSFADVNTKEAODLALRSAVEGA
      : * : . : :***:***: .** : * : * * * :***: * * . : * : : : :***
gi|reesei| VLLKNDGTLPLSKKVR--SIALIGPWANATTOMQGNYYGPAEYLISPLEAAKKAGYHVNEE
gi|fvert| VLLKNDGTLPLKLLKKKDSVAMIGFWANDTSKLQGGYSGRAFPLHSPLYAAEKLGLDTNVA
      ***** . * : * : * : * * * : : * : * * * : * * * : * * * : * .
gi|reesei| LGTEIAGNSTTG--FAKAIAAAKSDAIIYLCGIDNTIEQEGADRTDIAWPGNQDLIKQ
gi|fvert| WGPTLQNSSSHDNWTTNAVAAAKSDYILYFCGLDASAAGEDRDRENLDWPESQLTLLQK
      * , : .*: . : :*:***** * : * : * : * , * * : : * * . * * : :
gi|reesei| LSEVGKPLVVLQMGGGQVDSSSLKSNKKVNSLVWGGYPGQSGGVALFDILSGKRAPAGRL
gi|fvert| LSSLKPLVVIQLG-DQVDDTALLKNKKINSILWVNYPGQDGGTAVMDLLTGRKSPAGRL
      ** ,*****: * : * .*** : : * .***:***: * .*** ,** .*: : * : : : :***
gi|reesei| VTTQYPAEYVHQFPQNDMNLRPDGSNPGQTYIWYTGKPVYEEFGSGLEYTTFKETLASHP
gi|fvert| PVTQYPSKYTEQIGMTDMDLRPT-KSLPGRTYRWYS-TPVLPYGFGLHYTKFQAKFKSN-
      .*****: * .*: .**:* * * * : * * : * * . * * .*: : : * :
gi|reesei| KSLKENTSSILSAPHPGYTYSEQIPVFTFEANIKNSGKTESPYTAMLFVRTSNAGPAPYP
gi|fvert| -KLTFDIQKLLKG--CSAQYSDTCALPPIQVSVKNTGRITSDFVSLVFIKS-EVGPKPYP
      .* .*: . : * . . * : . : : : : * : : : : : : : * * *
gi|reesei| NKWLVGFDRLADIKPGHSSKLSIPIPVSALARVDSHG NRIVYPGKYELALNTDES VKLEF
gi|fvert| LKTLAAYGRLHDVAPSSTKDISLEWTLDNIAARRGENGLVVYPGTYTLLLEPTQAKIQV
      * * . . : * * : * . : : : : . : : * * . : * : : * * . : * : :
gi|reesei| ELVGEEVTIENWPLEEQQIKDATPDA
gi|fvert| TLTGKKAILDKWPQDPKSA-----

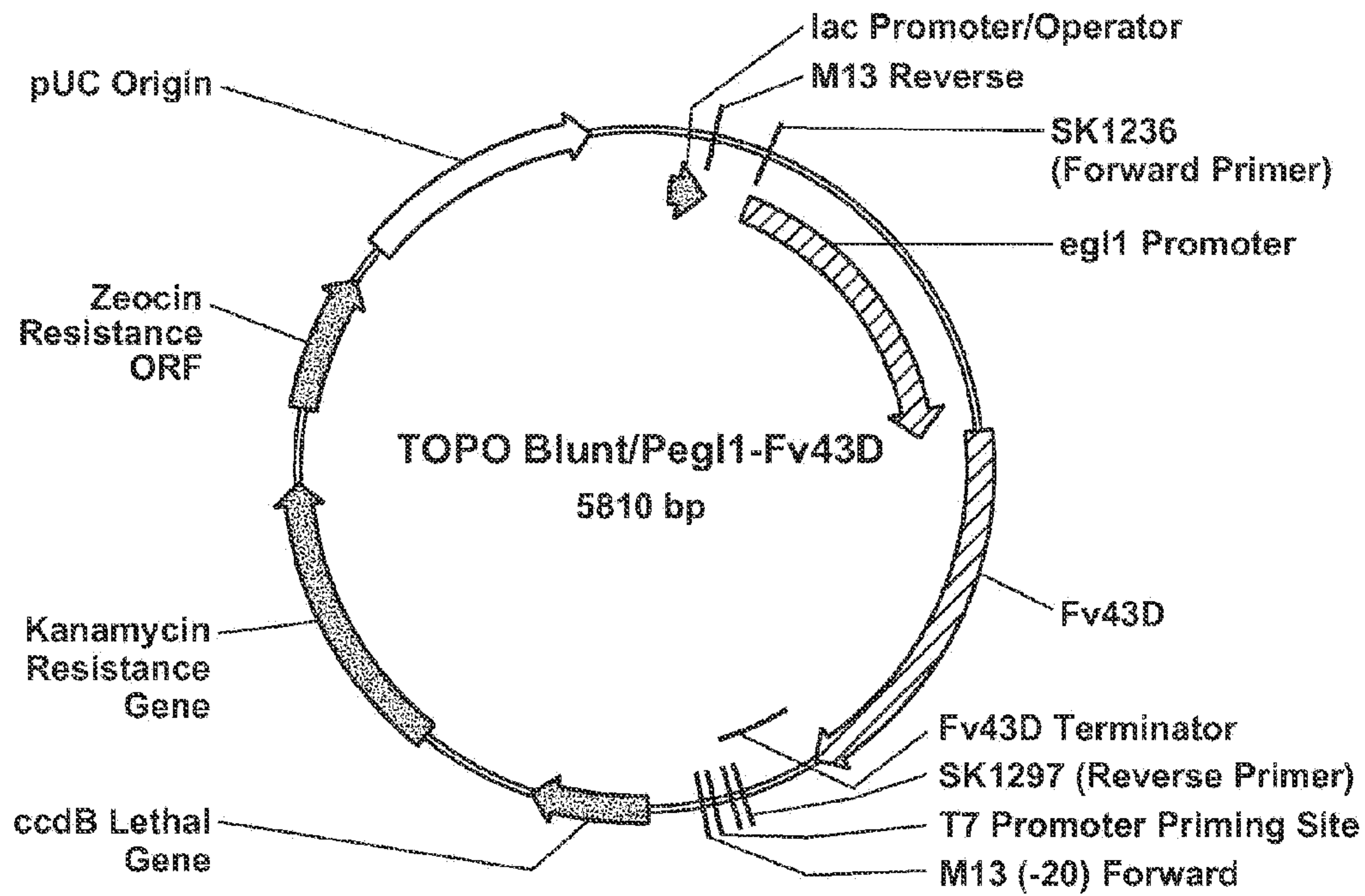
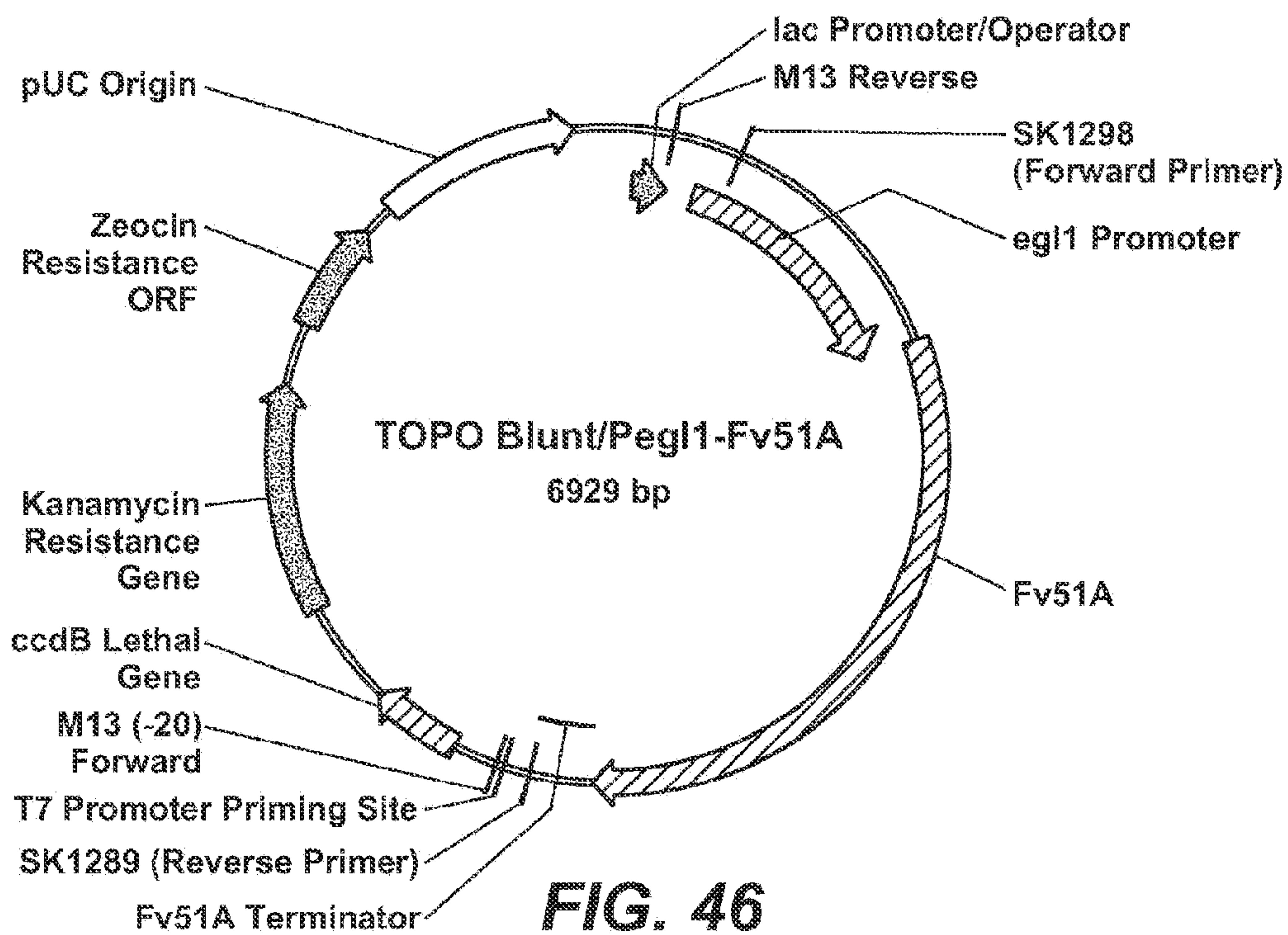
```

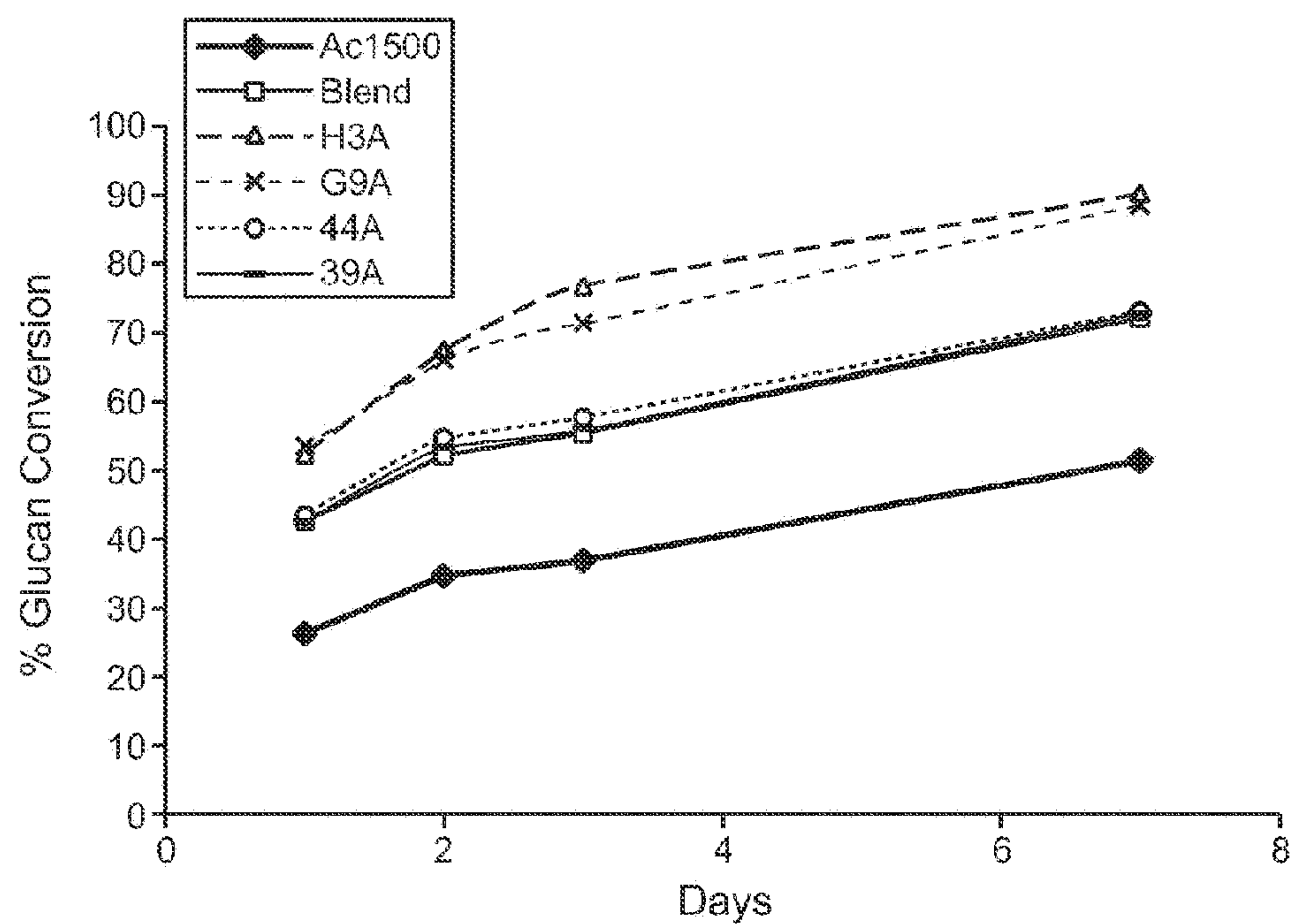
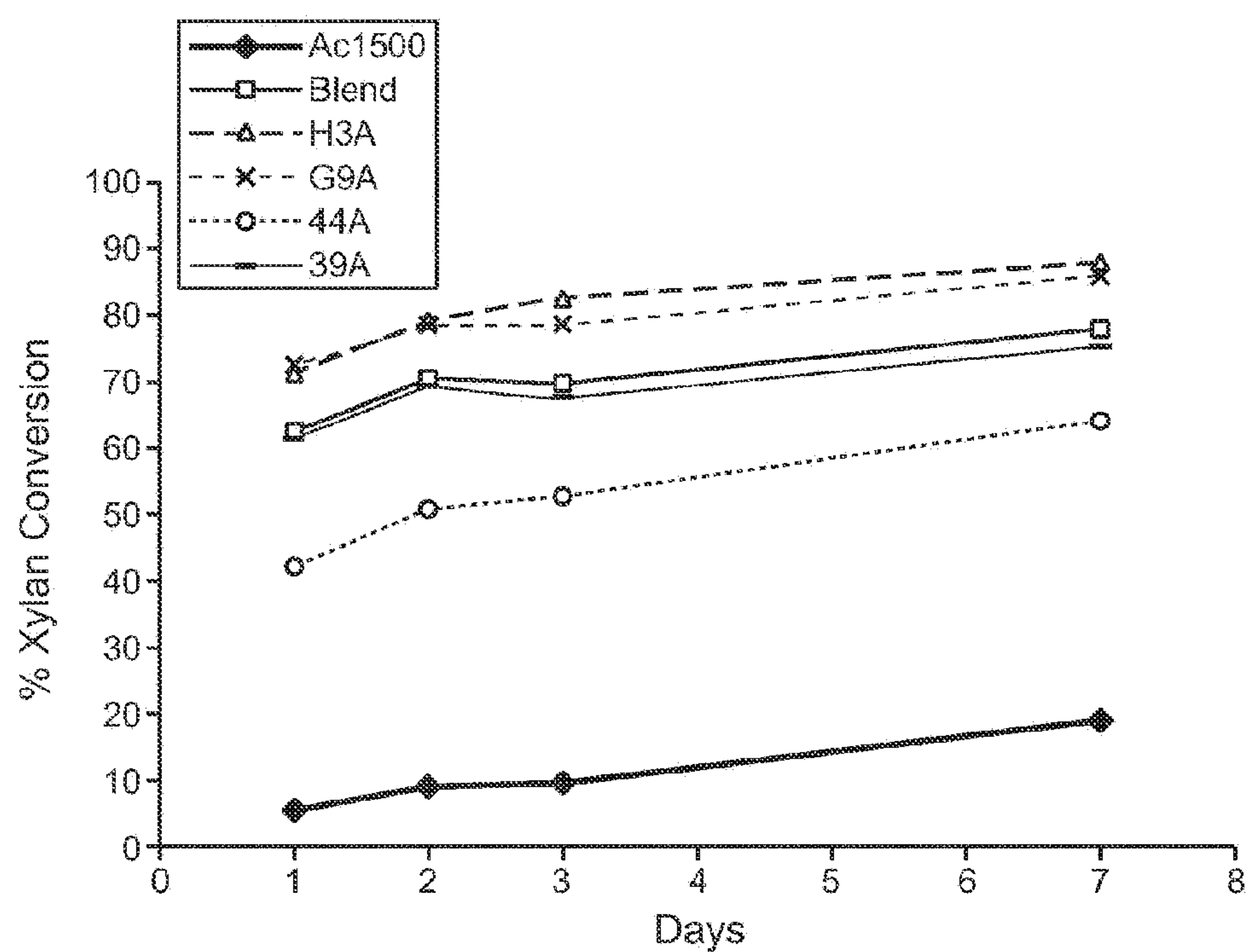
FIG. 38

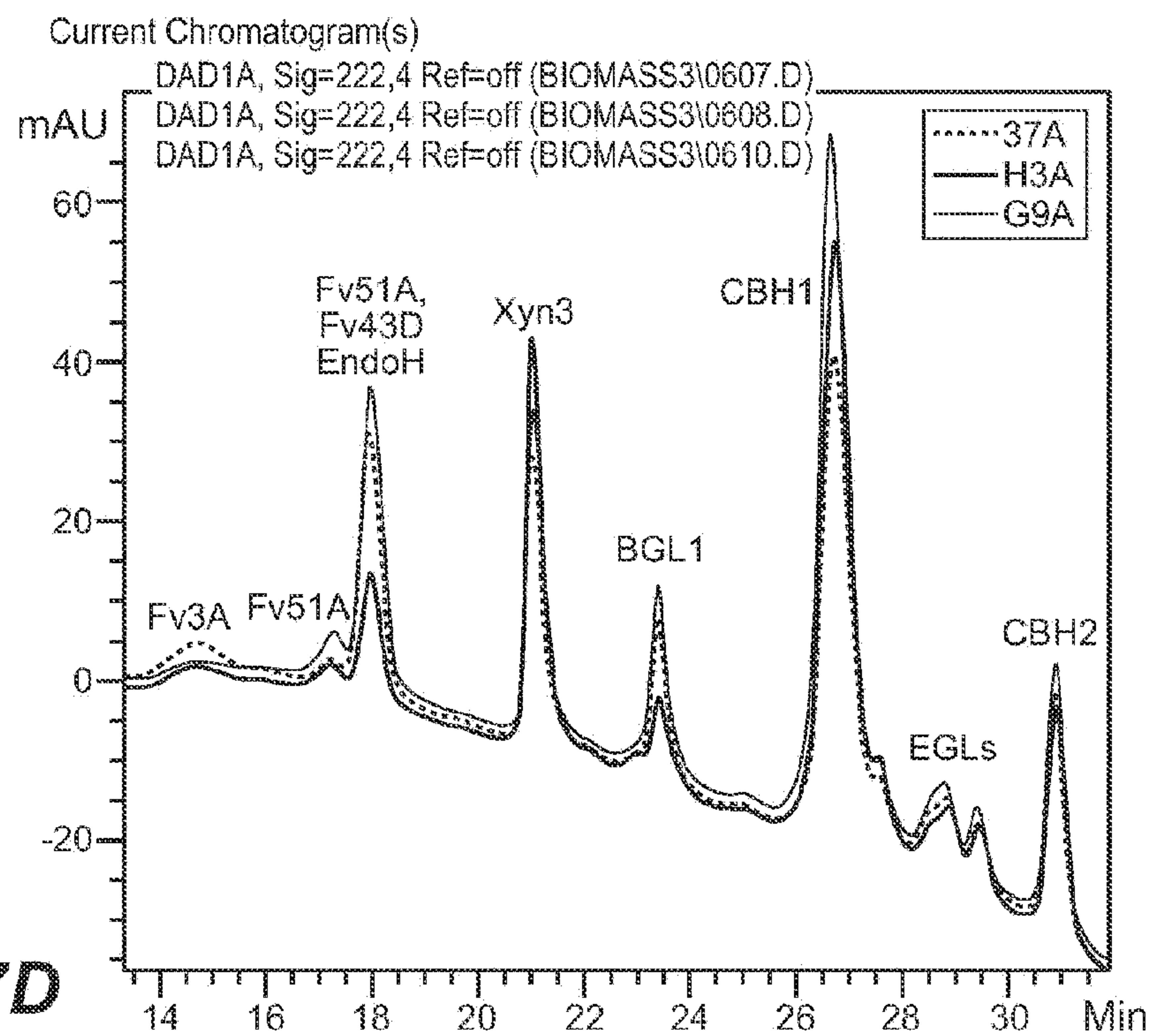
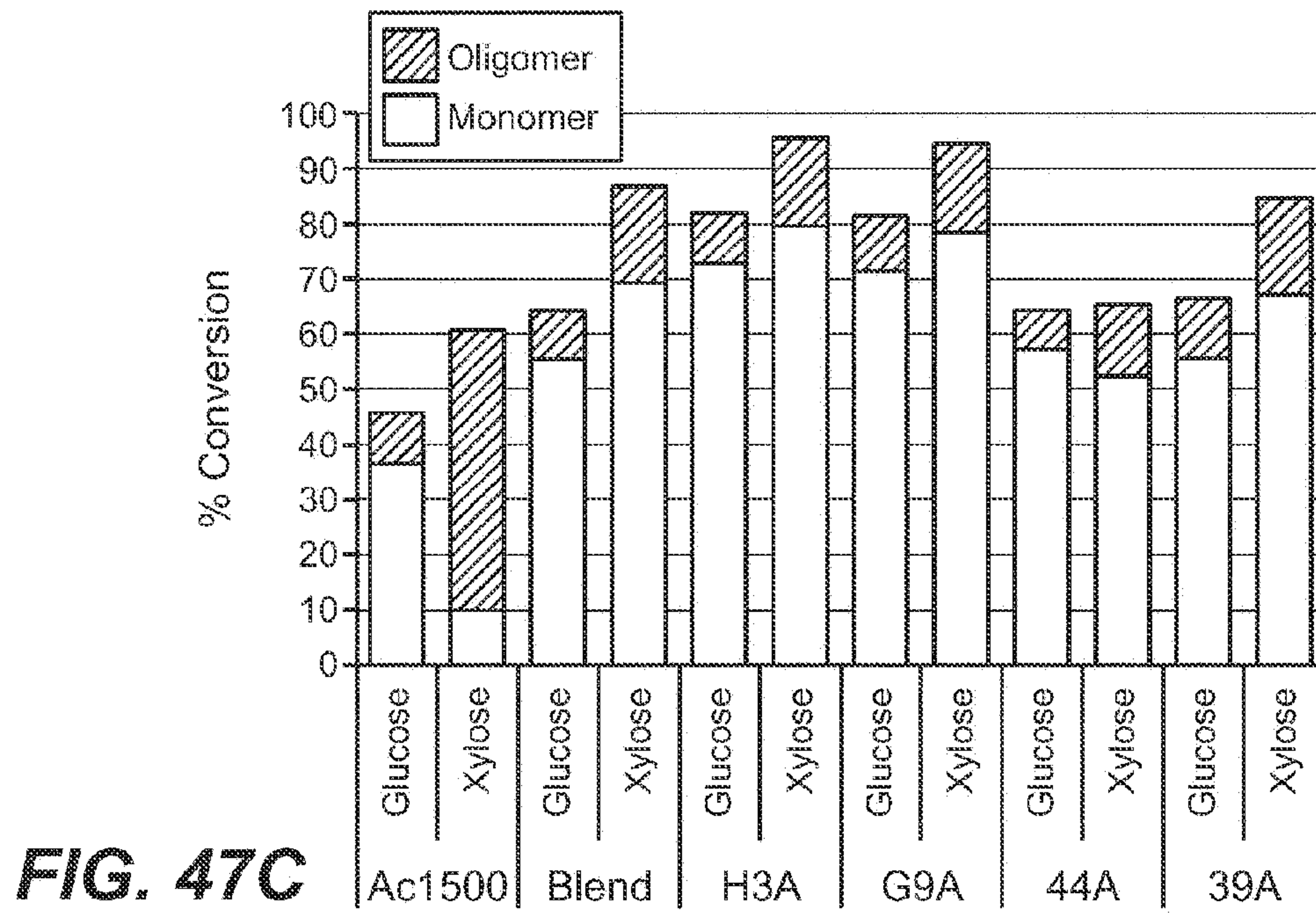


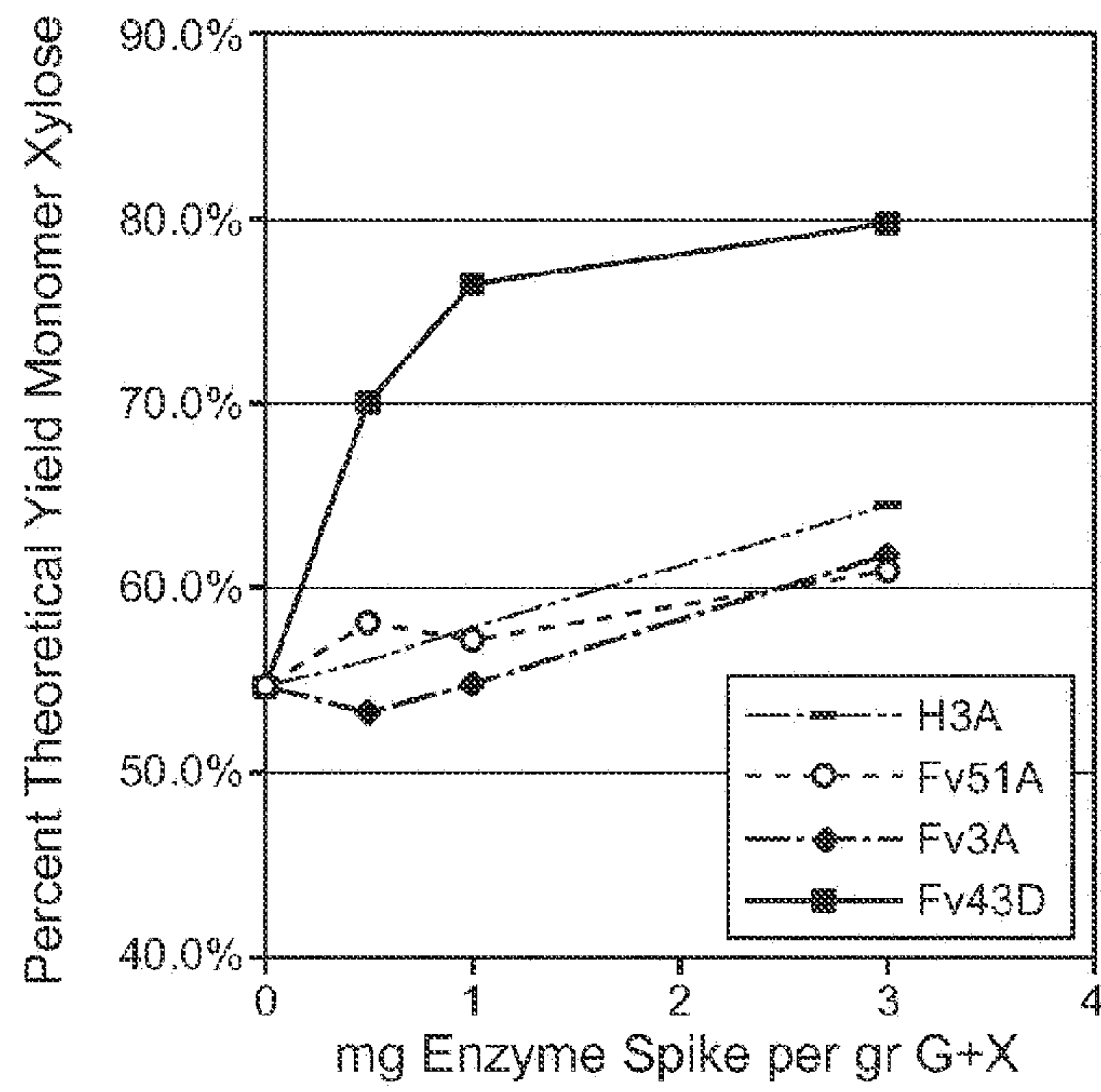
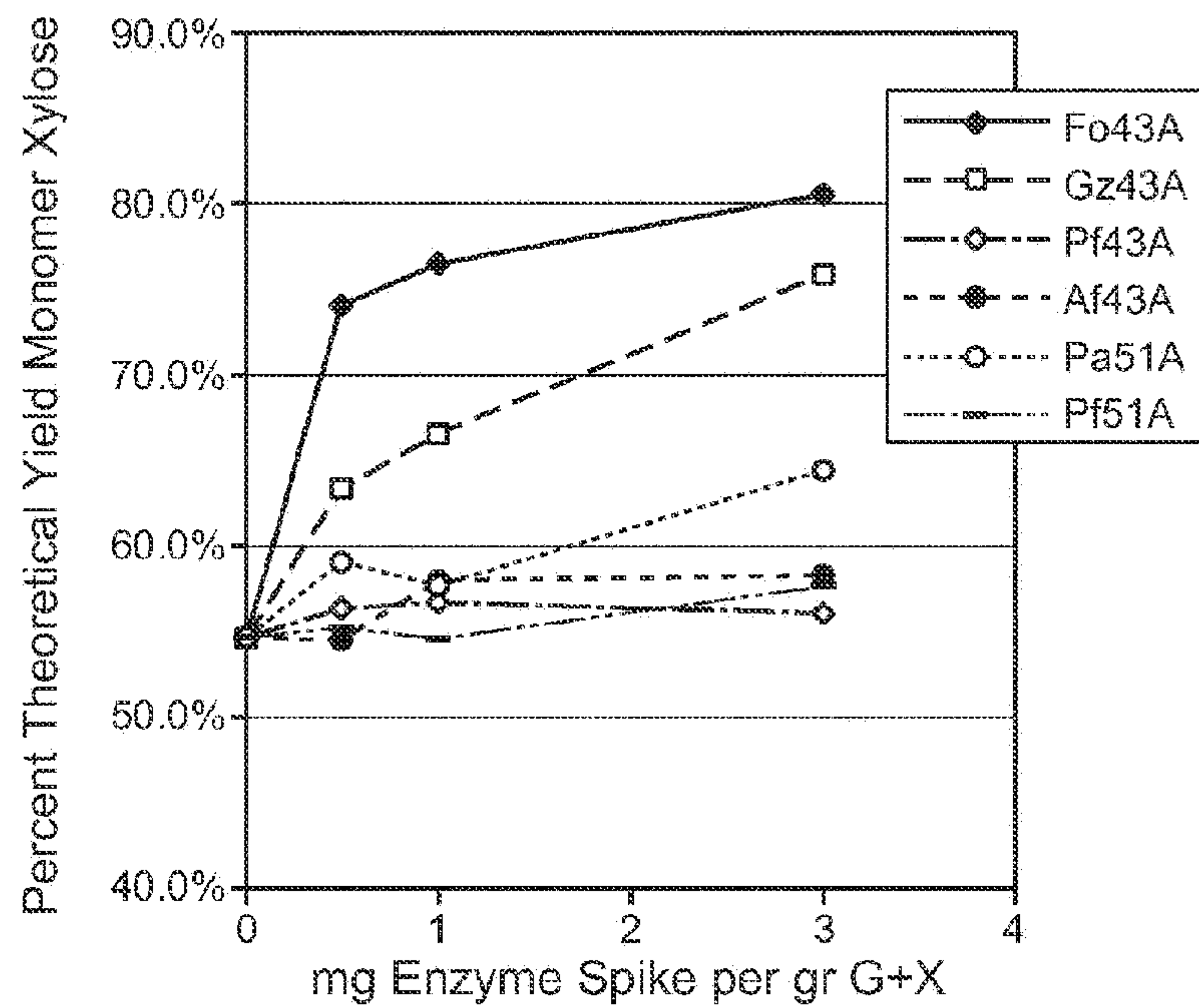


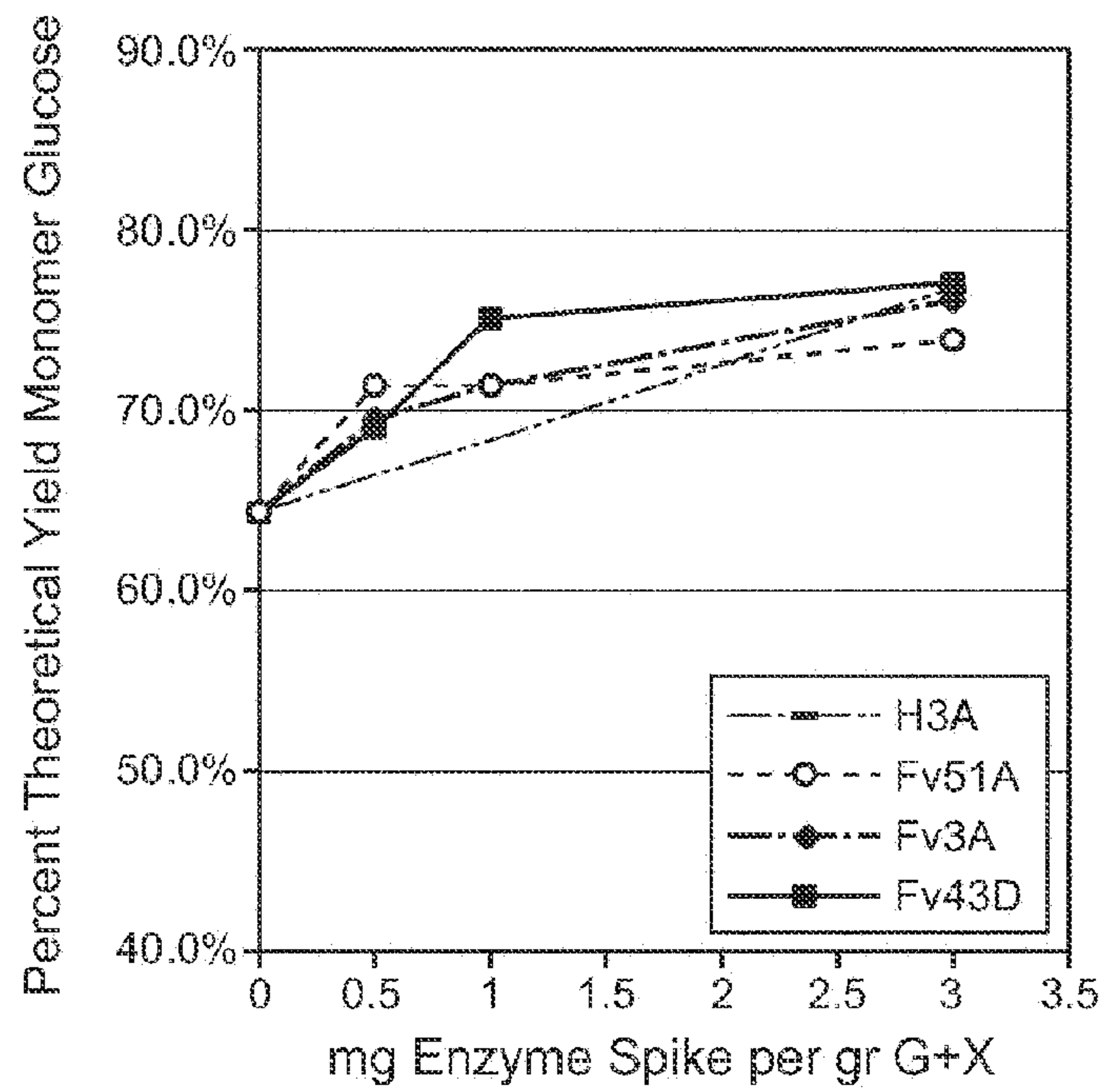
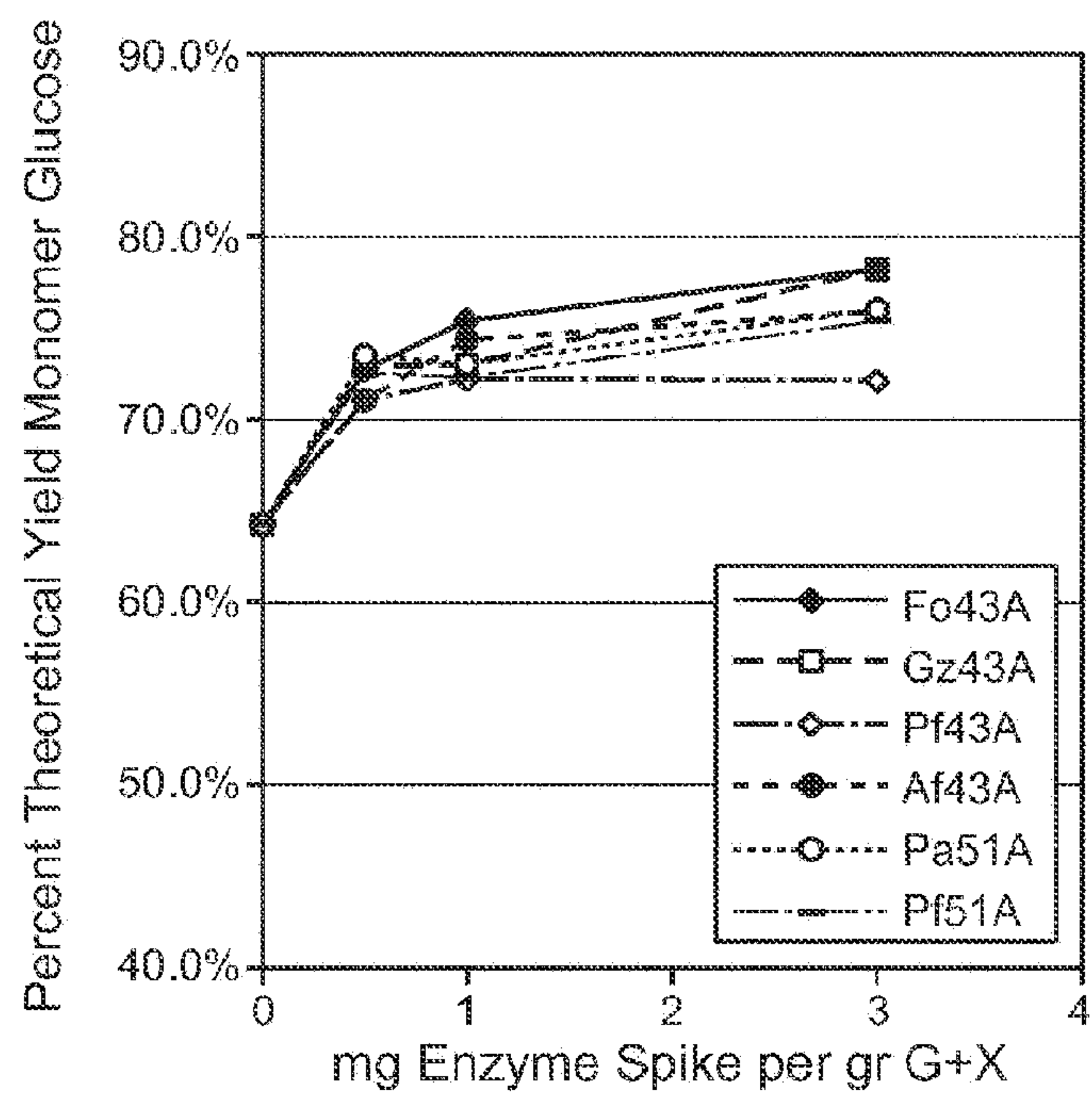
**FIG. 43****FIG. 44**

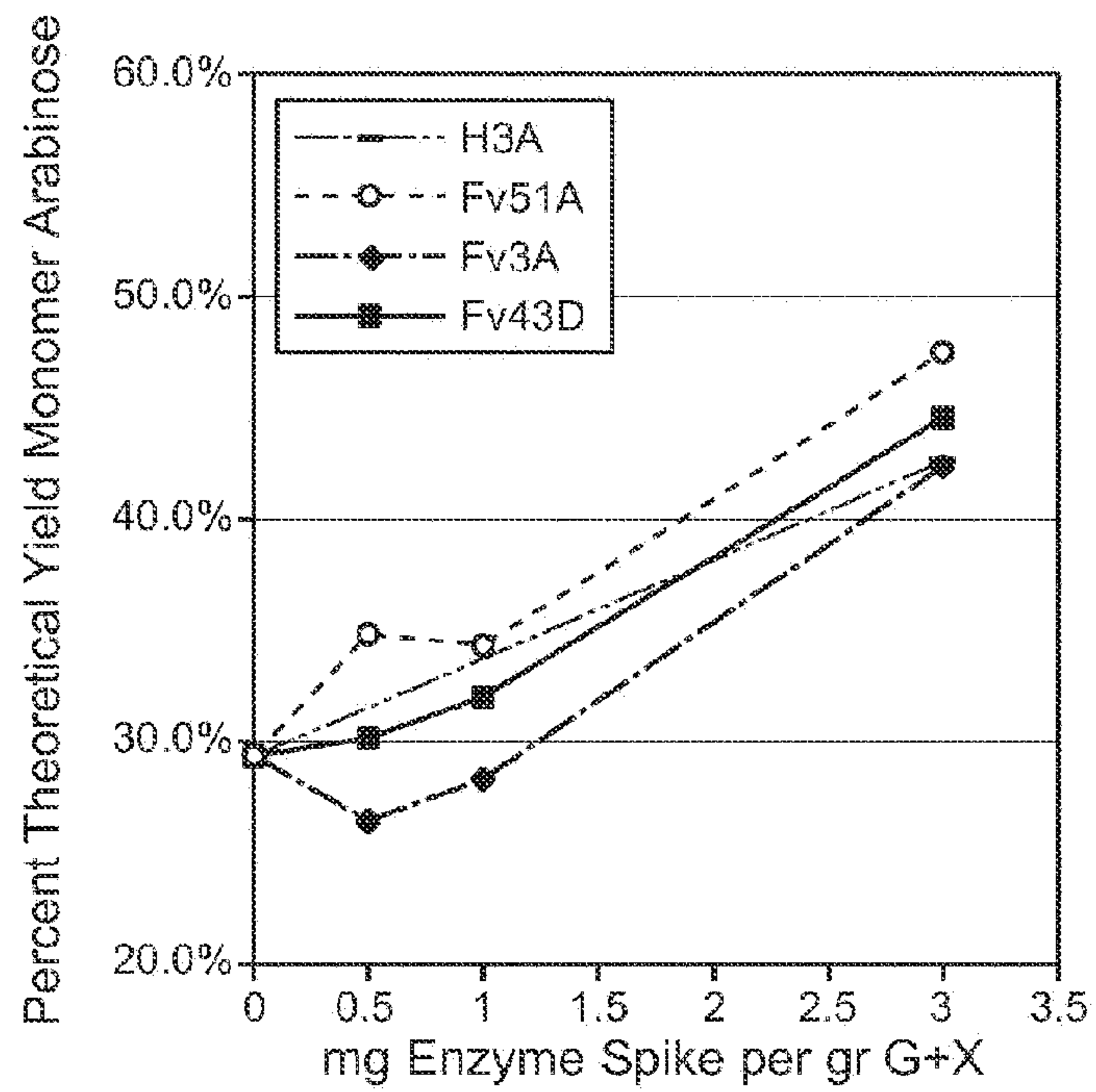
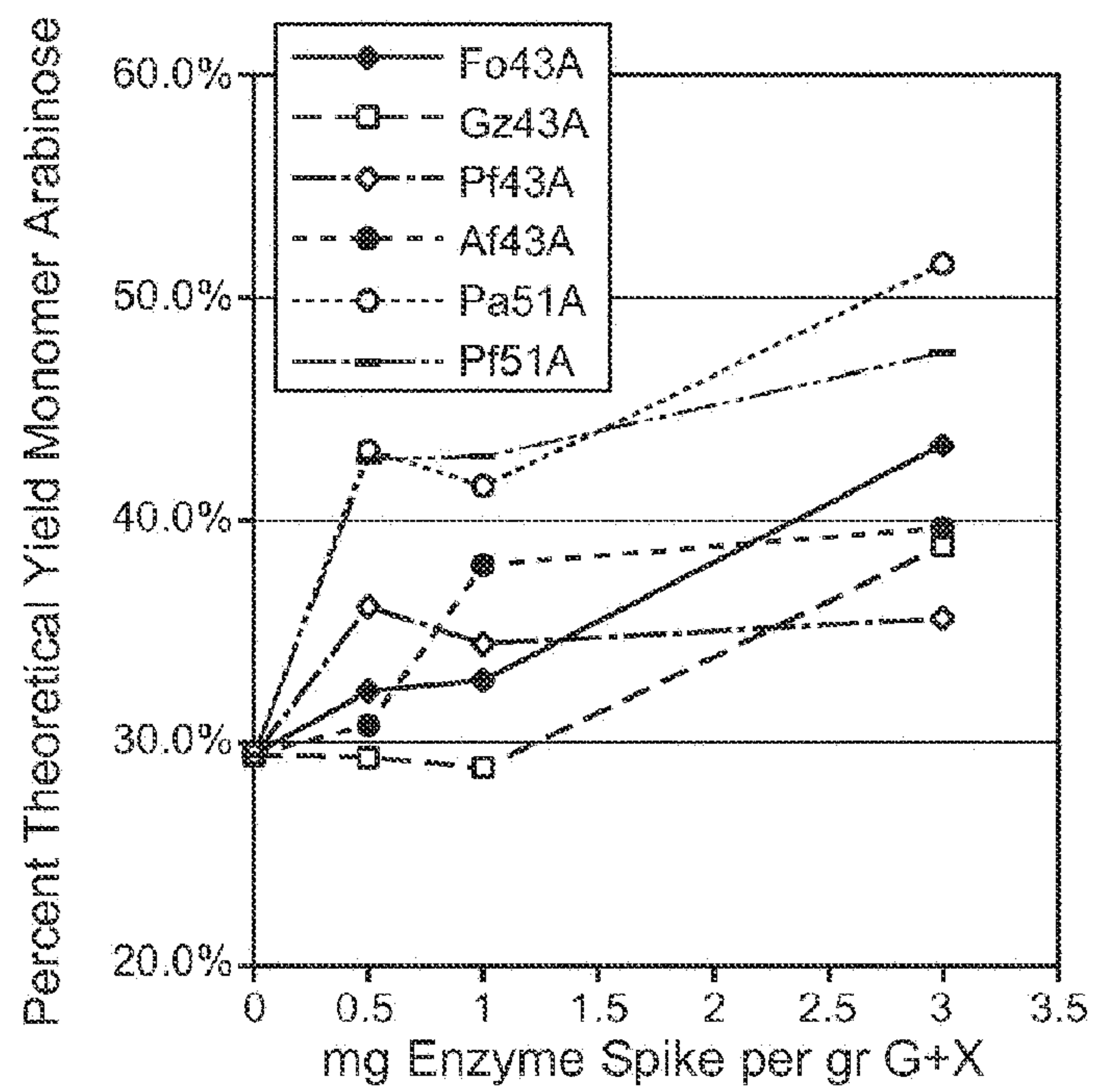
**FIG. 45****FIG. 46**

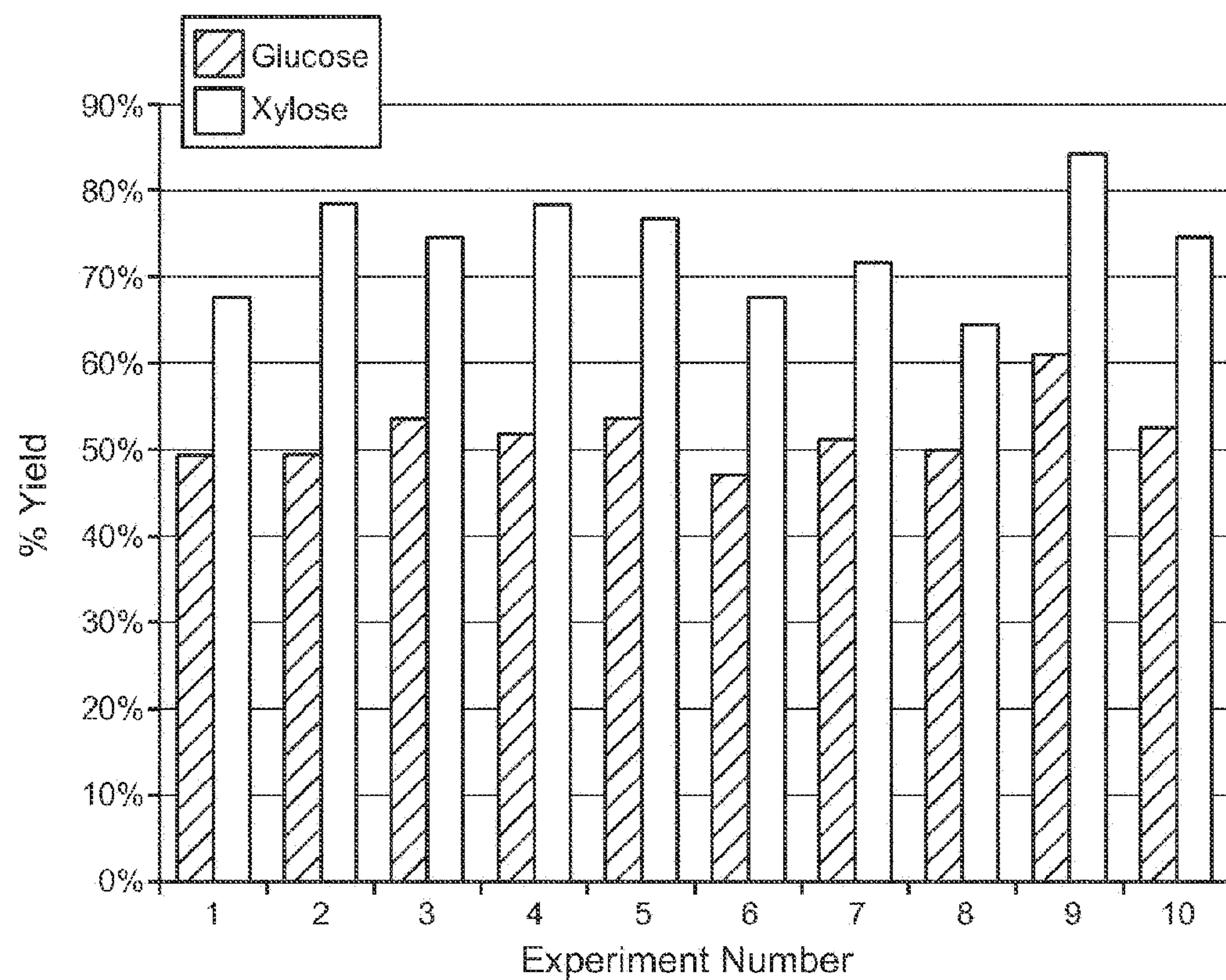
**FIG. 47A****FIG. 47B**



**FIG. 48A****FIG. 48B**

**FIG. 49A****FIG. 49B**

**FIG. 50A****FIG. 50B**

**FIG. 51**

Fv39A	:	MHYATLTTLVLALTNNVAAQQGTATVDLSKNHGPALGSGFIYGWPDNGTSVDT SIPDF	:	60
Q9ZFM2_XYN	:	-----MKVVNVPSNGREKFKKNW--KFCVGTGRL-GLALQKEYLDH--LKL	:	41
P36906_XYN	:	-----MIKVRVPDPSDKKFSDRW--RYCVGTGRL-GLALQKEYLEF--LKY	:	41
	:	* * * * *	:	
Fv39A	:	LVTDIKFNSNRGGGAQIPSLGWARGGYEG-----YLGR-FNSTLSNRYRTTRKYNA	:	109
Q9ZFM2_XYN	:	VQEKIGFRYIRGHGLLSDDVGIIYREVEIDGEMKPFYNYFTYIDRIVDSYLA--LNIRPF-I	:	98
P36906_XYN	:	VKENIDFKYIRGHGLLCDDVGIIYREDVVGDEVKPFYNYFTYIDRIFDSFLE--IGIRPF-V	:	98
	:	* * * * *	:	
Fv39A	:	DFILLPHDLWGADGG-----QGSNSPFFPGDNGNWTM-ELFWNQLVSDLKAHNMLEGLVID	:	164
Q9ZFM2_XYN	:	EFGFMPKALASGDQTVFYWKGNVTP-PKDYNNKWRDLIVAVVSHFIERYGIEEVRT-WLFE	:	156
P36906_XYN	:	EIGFMPKKLASGTQTVFYWEGNVTP-PKDYEKWSDLVKAVLHHEFISRYGIEEVVK-WPFE	:	156
	:	A * * * * *	:	
Fv39A	:	VWNEPDIDIFW-DRPWSQFLEYYNRAIKLLRKTLLSGPAMAHSP--LSDDKWHTWLQ	:	223
Q9ZFM2_XYN	:	VWNEPNLVNWKDANKQEFYFKLYEVTARAVKSVDPHLQVGGPAICGS---DEWITDFL	:	212
P36906_XYN	:	IWNEPNLKEFWKDADEKEYFKLYKVTAKAIKEVNENLVKVGGAICGA---DYWIEDFL	:	212
	:	* * * * *	:	
Fv39A	:	SVAGNKTVP-DIYSWHQIGAW-----EREPE-DSTIPDFTTLRA---QYGVF	:	264
Q9ZFM2_XYN	:	HFCAERRVPVDFVSRHAYTSKAPHKKTFEYYQELELEPPEDMLEQFKTVRALIRQSPFF	:	272
P36906_XYN	:	NFCYEENVPVDFVSRHATTSK-QGEYTPHLIYQ--EIMPSEYMLNEFKTVRE--IKNSHFF	:	269
	:	N * * * * *	:	
Fv39A	:	EKPIDVNEYAARDEQNPAANSVYYLSQLERHNLRLGLRANWGSGLHNMGNLIYSTGTS	:	324
Q9ZFM2_XYN	:	HLPLHITTYNT--SYSPINPVHDTALNAAAYIARILSEGGDYVDSFSYWTFSVDVFEEMDVP	:	330
P36906_XYN	:	NLPFHITTYNT--SYSPQNVPVHDTPFNAAYIARILSEGGDYVDSFSYWTFSVDVFEERDVP	:	327
	:	* * * * *	:	
Fv39A	:	EGTYY-----PNGEWQAQYKYAAMAGQRL-----VI	:	350
Q9ZFM2_XYN	:	KALFHGGEGGLVALHSIPKPTFHAFTEFNALGDELLYRDGEMIVTRRKDGSIAAVLWNLYM	:	390
P36906_XYN	:	RSQFHGGEGGLVALNMIPKPTFYTFKFFFNAMGEMLYRDEHMLVTRRDDGGSVALIWNNEVM	:	387
	:	* * * * *	:	
Fv39A	:	KASSDLKFQVVFATKQGRKIKIIAGTRTVQAKYNIKISGLEVAGLPKMGTVKVRTYRFDWA	:	410
Q9ZFM2_XYN	:	EKGECLTKEVQLVIPVSFSAVFIKRQIVNEQYCNAMRVWKOMGRPRFPSPQAVETLPS-A	:	449
P36906_XYN	:	DKTENPDEDYEVEIPVRFRDVFIKRQLIDEEHGNPWGTWIHMGRPRYPSKEQVNTLREVA	:	447
	:	* * * * *	:	
Fv39A	:	GP-----NGKVDGPVDLGEKKYT-YSANTVSSPST-----	:	439
Q9ZFM2_XYN	:	QPHVMTEQRRATDGVIIHLSIVLSKNEVTLLIEIEQVRDETSTYVGLDDGEITSYSS	:	504
P36906_XYN	:	KPEIMTSQPVANDGYLNLKFKLGKNAVVLVELTERIDESSTYIGLDDSKINGV---	:	500

FIG. 52

Fv43D ---MQLKFLSSALLSLTGNCAAQDTNDIPPLITDLWSADPSAHVFEGKLWVYPSHDIEA
Fo43A ---MQLKFLSSALLFSLTSKCAAQDTNDIPPLITDLWSADPSAHVFEGKLWVYPSHDIEA
Gz43A ---MKSLLFP--LLSFVG--QSLATNDDCPLITSRWTADDPSAHVFNDTLWLYP SHDIDA
Pf43A --MLQRFAYILPLALLSVG--VKADN-----PFVQSIYTADDPAPMVYNDRVYVFMHDNTG
Fv43A -MWLTSPLLFASTLLGLTGVALADN-----PIVQDIYTADDPAPMVYNGRVYLEFTGHDNDG
Fv43B --MRFSWLLCPLLAMGSALPETKTDVSTYTNPVLPGWHSDPSC-IQKDGLFLCVTSTFIS
Af43A -----MAAPSLSYPTGIQSYTNPLFPGWHSDPSCAYVAEQDTFFCVTSTFI
Pf43B ----MSRSILPYASVFALLGGAIAEP----FLVLNSDFPDPSLIETSSGYAFGTTGNGV
Fv43E MKVYWLVAWATSLTPALAGLIGHRRATTFNNP IISDFPDNDVFLGPDNYYYFSASNFHE

Fv43D NVVNGTGGAQYAMRDYHTYSMKTIYGKDPVIDHGVALSVDDVPWAKQQMWAPDAAYK--N
Fo43A NVVNGTGGAQYAMRDYHTYSMKSTYKDPVVDHGVALSVDDVPWAKQQMWAPDAAHK--N
Gz43A GFENPDGGAQYAMRDYHVYSIDKIYGSLP-VDHGVALSVDDVPWASRQMWAPDAAHK--N
Pf43A -----ATYINMTDWHLFSSADMANWQD---HGIPMSLANFTWANANAWAPQVIPR--N
Fv43A -----STDFNMTDWRLFSSADMVNWQH---HGVPMSLKTFSWANSRAWAGQVVAR--N
Fv43B FP----GLPVYASRDLVNWRRLISHVWNRE---KQLPGISWKTAGQQQGMYP TIRYH--K
Af43A AFP----CLPLYASRDLQNWKLASNIFNRP---SQIPDLR-VTDCQQSCIYAPT LRYH---E
Pf43B N-----AQVASSPDFNTWTLLSGT-----DALPGFPFSWVASSPQIWAPDVLVKA-D
Fv43E SP----GAPVLKSKDLLNWDLIGH SIPRLNFGDGYDLPPGSRXYRG-GTWASSLR YRKS N

Fv43D GKYYLYFPAK-DK-DEIFRIGVAVSNKPSGPFK---ADK-SWIPGTYSIDPASYVDTNGE
Fo43A GKYYLYFPAK-DK-DEIFRIGVAVSNKPSGPFK---ADK-SWIPGTYSIDPASYVDTDNE
Gz43A GKYYLYFPAK-DK-DDIFRIGVAVSPTPGGPFV---PDK-SWIPHTFSIDPASFVDDDDR
Pf43A GQFYFYAPVR-HN-DGSMALGVGVSSITGPHYH---DAIGKPLVENNEIDPTVFIDDDGQ
Fv43A GKFYFYVPVRNAK-TGGMALGVGVSTNILGPYT---DALGKPLVENNEIDPTVYIDTDGQ
Fv43B GTYYVICEYLGVG-DLIGVIFKTTNPWDESSWS---DFV---TFKPNHIDPDLFWD DDGK
Af43A GQFYLI VSYLGP--QTKGLLFTSSDPYDDA AWS---DPL---EFAVHG IDPDIFWDH DGT
Pf43B GTYVMYFSASAASDSGKHCVGAATATSPEGPYTPVDSAVACPLDQGGAIDANGFIDTDGT
Fv43E GQWYWIGCIN-----FWQTWVYTASSPEGPWY-----NKG NFGDNNCYDNGILIDDD ET

Fv43D AYLIWGGI-WGGQLQAWQDHKTFN ESWLGDKAAPNGTNALSPQIAKLSKDMHKITETPRD
Fo43A AYLIWGGI-WGGQLQAWQDKKNFNESWIGDKAAPNGTNALSPQIAKLSKDMHKITETPRD
Gz43A AYLAWGGI-MGGQLQRWQDKNKYNES---GTEPG-NGTAAALSPQIAKLSKDMHTLA EKPRD
Pf43A AYLYWG-----NFDLWYVKLNQDMISYSGSPTQ
Fv43A AYLYWG-----NFGLYYVKLNQDMLSYSGSINK
Fv43B VYCATHG----ITLQEI DLETGELSPELNIWNGTGGVWPEGPHIYKRDGYYYLMIAEGGT
Af43A VYV TSAED-QMIKQYTLDLKTGAIGPVDYLWNGTGGVWPEGPHIYKRDGYYYLMIAEGGT
Pf43B IYVVYKID-----GNSLDGDGTTHPTPIMLQQMEADGT
Fv43E MYVVYGSGEVKVSQLSQDGF SQVKSQVVFKN TDIGVQDBLEGNRM YKING-----LYYI

FIG. 53A

Fv43D LVILAPETGKPLQAEDNKRRFFEGP-----WVHKRGKLYYLMYSTG-----
Fo43A LVILAPETGKPLQAEDNKRRFFEGP-----WIHKRGKLYYLMYSTG-----
Gz43A MLILDPKTGKPLLSEDEDRRFFEGP-----WIHKRNKIYYLTYSTG-----
Pf43A IPLTTAGFGTRTGNAQRPTTFFEAP-----WVYKRNGIYYIAYAAD-----
Fv43A VSLTTAGFGSRPNNAQRPTTFFEGP-----WLYKRGNLYYMIYAAN-----
Fv43B ----AEDHAITIAARARKITGPYEAYNNNPILTNRGTSEYFQTVGHGDLFQDTKGNWWGLC
Af43A ----ELGHSETMARSRTGTGPWEPYPHNPLLSNKGTSEYFQTVGHADLFQDGNGNWWAVA
Pf43B ---TPTGSPITQLIDRSDLDGPLIEAP-----SLLLSNGIYYLSFSSN-----
Fv43E LNDSPSGSQTWIWKSKSPWGPYESKVLADKVTPPISGGNSPHQGS LIKTPNGGWY-----

Fv43D -DTHFLVYATSKN---IYGPYT-----YQGKILDPVDG-----WTTHG
Fo43A -DTHFLVYATSKN---IYGPYT-----YRGKILDPVDG-----WTTHG
Gz43A -TTHYLVYATSKT---PYGPYT-----YQGRILEPVDG-----WTTHS
Pf43A CCSEDIRYSTGTS---ATGPWT-----YRGVIMPTQGSS-----FTNHE
Fv43A CCSEDIRYSTGPS---ATGPWT-----YRGVVMNKAGRS-----FTNHP
Fv43B LATRITAQGVSPMGREAVLFNGTWNKGEWPKLQPVGRMPGNLLPKPTRN-----VPGD
Af43A LSTRSCPAWKNYPMCRETVLAPAAWEKCEWPVIQPVRCQMOC-PFPPPNKR----VPRCE
Pf43B YYNTNYYDTSYAYASSITGPWT-----KQSAPYAPLLVTGT-----ETSND
Fv43E FMSFTWAYPAGRLPV LAPITWG-----SDGEPIILVKGANGGWGSSYPTLPGT

Fv43D SIVEYKGQWWLFFAD-AHTSGKDYL RQVKARKI WYDKDG-----KILLTRPKI-----
Fo43A SIVEYKGQWWLFFAD-AHTSGKDYL RQVKARKI WYDKNG-----KILLHRP-----
Gz43A SIVKYQGQWWLFYHD-AKTS GKDYL RQVKAKKI WYDSKG-----KILTKKP-----
Pf43A GIIDFQNN SYFFYHNGALPGGGGYQRSVCVEQFKYNADG-----TIPTIEMTTAG-----
Fv43A GIIDFENNSYFFYHNGALDGGSGYTRSVAVESFKYGS DG-----LIPEIKMTTQG-----
Fv43B GPFNADPDNYNLKKTKKIPPHFVHHRVPRDGA FSLSSKG---LHIVPSRNNVTG SVLPG
Af43A GGWIKQPDKVDFRPGSKI PAHFQYWRYPKTEDFTVSPRGHPNTLRLTPSFYNLTG-----
Pf43B GALSAPGGADFSVDG TKMLFHANLNGQDISGGRALFAAS-----ITEASDVVTLQ-----
Fv43E DGVTKNWTRTDTFRGTSLAPSW EWNHNP DVNSFTVNNGLT LRTASITKDIYQARN-----

Fv43D -----
Fo43A -----
Gz43A -----
Pf43A -----PAQIGTLN PYVRQEAETA AWSSGITEVCSEGGIDVGEINNG
Fv43A -----PAQLKSLN PYVKQEAETIAWSEGIETEVCSEGG LNVA FIDNG
Fv43B DEIELSGQRGLAFIGRRQTHTLFKYSVDIDFKPKSDDQEAGITV FRTQFDHIDL GIVRLP
Af43A -TADFKPDDGLSLVMRKQTDTLFTYTV DVSFDPKVADEEAGVTVFLTQQQHIDL GIVLLQ
Pf43B -----
Fv43E -----TLSHRTHGDHPTGIVKIDFSPMKDGDRA GLSAFRDQSAYIGIHRDNCK

FIG. 53B

```

Fv43D -----
Fo43A -----
Gz43A -----
Pf43A DYIK-----VKGVAFGS-GAHSFSARVASANSGGTIAIHLGSTTGTLVGTCTV
Fv43A DYIK-----VKGVDGSGTGAKTFSARVASNSGGKIELRLGSKTGKLVGTCTV
Fv43B TNQGSNKKSKLAFRFRATGAQNVPAK---VVPVPGWEKGVLSLHIEAANATHYNLGAS
Af43A TTEG-----LSLSFRFRVEGRGNYEGPLPEATVPVPKEWCGQTERLEIQAVSDTEYVFAAA
Pf43B -----
Fv43E FTIAT----KHGMNMDewNGTTTDLGQIKATANVPSGRTKIWLRLQLDTNPAGTGNTIFS

Fv43D -----
Fo43A -----
Gz43A -----
Pf43A PSTGGWQTWTTVTCSVSGASGTQ-----DVYFVFGGSGTGYLEN-----FDYWQFA
Fv43A TTTGNWQTYKTVDCPVSGATGTS-----DLFFVFTGSGSGSLFN-----FNWWQFS
Fv43B --SHRGKTLDIATASASLVSGGTGSFVGSLGPHYATCNGKSGVECPKGGDVYVTQWTYK
Af43A PARHQAQRQIISRANSLIVSGDTGRFTGSLVGVYATSNG-GAGSTP-----AYISRWRYE
Pf43B -----
Fv43E YSWDGVKYETLGPNFKLYNG-----WAFPIAYRFGIFNFAETALGGSIKVESFT

Fv43D -----
Fo43A -----
Gz43A -----
Pf43A -----
Fv43A -----
Fv43B PVAQEIDHGVFVKSEL
Af43A GRGQMIDFGRVVPsy-
Pf43B -----
Fv43E AA-----

```

FIG. 53C

Pa51A MIHLKPALAALLALSTQCVAIDLFVKSSGGNKTDDIMYGLMHEDINNSGDGGGIYAELISN
Fv51A MVRFSSILAAAACF-VAVESVNIKVDSEKGGNATSGHQYGFLLHEDINNSGDGGGIYAELIRN
Pf51A MGKMWESILVVLGLLSVGHATTINVSQSGGNKTSPLQYGLMFEDINHGGDGGGLYAELVRN

Pa51A RAFQGSEKFPSNLDNWSPVGGATLTLLQKLAKPLSSALPYSVNVANPKEGKGGKDKTKGKK
Fv51A RAFQYSKKYPVSLSGWRPINDAKLSLNRLDTPLSDALPVSMNVK---FGKGGK-----AKE
Pf51A RAFQGSTVYPANLDGYDSVNGAILALQNLTNPLSPSPSSLNVA-----KGS-----NNGS

Pa51A VGLANAGFWGMDVKRQKYTGSEFHVITGEYXGDFEVSLSRAITGETFGKKVVKGGSKKGGKWT
Fv51A IGFLNEGYWGMDVKKQKYTGSEFWVKGAYXGHFTASLSRNLTDDEVFGSVKVKSKANKKQWV
Pf51A IGFANEGWWGIEVKPQRYAGSFYVQGDYQGDFFDISLQSKLTQEVFATAKVRSSGKHEDWV

Pa51A EKEFELVPFKDAPNSNNTFVVQWDAEGAKDGSLDNLNLSLFPPTFKGRKNGLRIDLQAQTM
Fv51A EHEFVLTPNKNAPNSNNTFAITYDFPGA-DGALDFNLISLFPPTYKGRKNGLRVDLAEAL
Pf51A QYKYELVPKKAASNTNNTLTITTFDSKGLKDGSLNFNLISLFPPTYKNRPNGLRIDLVEAM

Pa51A VELKPTFLRFPGGNMLEEGNLTLDTWKWKYETIGPLKDRPGMAGVWEYQQTLGLGLVEYMEW
Fv51A EGLHPSLLRFPGGNMLEEGNLTNKTWWDWKDTLGPLNRPGFEGVWNYQQTHGLGILEEYLQW
Pf51A AELEGKFLRFPGGSDVEGVQAPYWKWNETVGDLDKDRYSRPSAWTYEESNGIGLIEEYMNW

Pa51A ADDMNLEPIVGVFAGLALDGSFVPESEMGWVIQQALDEIEEFLTGDAKTTKWGAVRAKLGH
Fv51A AEDMNLEIIVGVYAGLSLDGSVTPKDQLQPLIDDALDEIEEFIRG-PVTSKWGKKRAELGH
Pf51A CDDMGLEPILAVWDGHYLSNEVISENDLQPYIDDTLNQLEEFLMG-APDTPYGSWRASLGY

Pa51A PKPWKVKWVEIGNEDWLAGRPAGFESYINRFPMMMKAFNEKYPDIKIIASPSIFD----
Fv51A PKPFRLSYVEVGNEEDWLAGYPTGWNYSKEYRFPFMFLAIIKKAHPDLTVISSGASIDPVGK
Pf51A PKPWTINYVEIGNEDONLYG---GLETYIAYRFQAYYDAITAKYPHMTVMESLTEMPG---

Pa51A ---NMTIPAGAAGDHHPYLFPDEFVERFAKFDNLKDNVTLIGEAASTHPNG---GIAWE
Fv51A KDAGFDIPAPGIGDYHPYREPDVLVEEFNLFDDNNKYG--HIIGEVASTHPNG---GTGWS
Pf51A -----PAAAASDYHQYSTPDGFSQFNYPDQMPVTNRTLNGEIA TVYPNNPSNSVAVG

Pa51A GDLMPPLPWWGGSVAEAI FLISTERNNGDKIIIGATYAPGLRSLDRWQWSMTWVQHAADPALT
Fv51A GNLMPPYWWISGVGEAVALCGYERNADRI PGTFYAPILKNENRWQWAI TMIQFAADSAMT
Pf51A SPFPLYPWWIGSVAEAVFLIGEERNNSPKIIIGASYAPMFRNINNWQWSPTLIAFDADSSRT

Pa51A TRSTSWYVWRILAHHI IRETLPVDAPAGKPNFDPLFYVAGKSES-GTGIFKAAVYNSTES
Fv51A TRSTSWYVWSLFAGHPMTHTLPTTA-----DFDPLYVAGKNEDKGTLLIWKGAAAYNTTKG
Pf51A SRSTSWHVIKLLSTNKITQNLPTTWSGG--DIGPLYWVAGRNDNTGSNIFKAAVYNSTSD

Pa51A --IPVSLKFDGLNEGAVANLTVLTGPE-DPYGYNDPFTGINVVKEKTTFIKAGKGGKFTF
Fv51A ADVPVSLSFKGVKPGAQAE LTLTNKEKDPFAFNDPHKGNVVDTKKTVLKADGKGAFNF
Pf51A --VPVTVQFAGCN-AKSANLTI LSSDD--PNASNYPG-GPEVVKTEIQSVTANAHGAFEF

Pa51A TLPGLSVAVLETADAVKGGKGGKGGKGGKGN
Fv51A KLPNLSVAVLET LK-----KGGPYSS
Pf51A SLPNLSVAVLKTE-----

FIG. 54

	*	20	*	40	*	
xyn3	:	-MKANVILC--LLAPLVAALPTE	TIHLDP	ELAALRANLT	ERTADLWDRQA	: 47
P56588	:	-----			QA	: 2
P23360	:	MVRPTILLTSLLLAPFAAASP	-----		LEERQA	: 28

		60	*	80	*	100	
xyn3	:	SQSIDQLIKRK	GKLYFGTATDRG	LLQRE--KNA	AI IQADLGQV	TIPENS	SMKW : 96
P56588	:	SVSIDAKFKAHG	KKYLGTIGDQY	TLTKNTKNPA	IIKADFGQLT	PENS	SMKW : 52
P23360	:	AQSV DQLIKARG	KVYFGVATDQNR	LTG--KNA	AI IQADFGQV	TIPENS	SMKW : 77

	*	120	*	140	*	
xyn3	:	QSLENNQGQLN	WGDADYLVNFA	QQNGKSIRGH	TLIWHSQLFA	VNNINNA : 146
P56588	:	DATEPNRGQFT	FSGSDYLVNFA	QSNGKLIRGH	TLVWHSQLP	GWVSSITDK : 102
P23360	:	DATEPSQGNFN	FAGADYLVNWA	QQNGKLIRGH	TLVWHSQLP	SWVSSITDK : 127

		160	*	180	*	200	
xyn3	:	DTLRQVIRTHV	STVVGRYKGI	RAWDVVNEIF	NEDGTLRSSV	FSRLLGEE : 196	
P56588	:	NTLISVLKNH	TTVMTRYKGI	YAWDVLNEIF	NEDGSLRNSV	FYNVIGED : 152	
P23360	:	NTLTNVMKNH	TTLMTRYKGI	RAWDVVNEAF	NEDGSLRQIV	FLNVIGED : 177	

	*	220	*	240	*	
xyn3	:	FVSI AFRAARD	ADPSARLYIN	DYNLDRANYG	KVNG LKTYV	SKWISQGVPI : 246
P56588	:	YVRI AFETARS	VDPNAKLYIN	DYNLDSAGYS	KVNG MVSHV	KKWLAAGIPI : 202
P23360	:	YIPI AFQTARA	ADPNKLYIN	DYNLDSASYP	KTQAIVNRV	KQWRAAGVPI : 227

		260	*	280	N	*	300	
xyn3	:	DGIGSQSHLSG	GGGSGTLGAL	QQLATVPVTE	LAIT <u>EL</u> DIQ	GAPTIDYTQV : 296		
P56588	:	DGIGSQTHLG	AGAGSAVAGA	LNALASAGT	KEIAIT <u>EL</u> DI	AGASSIDYVNV : 252		
P23360	:	DGIGSQTHLS	AGQGAGVLQ	ALPLLASAGT	PEVAIT <u>EL</u> D	VAGASPIDYVNV : 277		

	*	320	*	340	*	
xyn3	:	VQACLSVSKCV	GITVWGISDK	DSWRAS	TNPLLEDAN	FNPKPAYNSIVGIL : 346
P56588	:	VNACLNQAKCV	GITVWGVADP	DSWRSSSP	LLFDGNYNP	KAAAYNAIANAL : 302
P23360	:	VNACLVQSCV	GITVWGVADP	DSWRAS	TTPLLEDGN	FNPKPAYNAIVQDL : 327

xyn3	:	Q--	: 347
P56588	:	--	: -
P23360	:	QQ	: 329

FIG. 55A

1	M V S F S Y L L L A C S A I G - A L A A P V E P E T T S F N	AfuXyn2
1	M I S I S S L S F G L A A I A G A Y A L P S D - K S V S L A	AfuXyn5
1	M V S F T S L L A A S P P S R A S C R P A A E V E S V A V E	Xyn2
30	E T A L H E F A E R A G T P S S T G W N N G Y Y Y S F W T D	AfuXyn2
30	E R - - - - - Q T I T T S Q T G T N N G Y Y Y S F W T N	AfuXyn5
31	K R - - - - - Q T I Q P G - T G Y N N G Y F Y S Y W N D	Xyn2
60	G G G D V T Y T N G A G G S Y S V N W R N - - V G N F V G G	AfuXyn2
53	G A G S V Q Y T N G A G G E Y S V T W A N Q N G G D F T C G	AfuXyn5
53	G H G G V T Y T N G P G G Q F S V N W S N - - S G N F V G G	Xyn2
88	K G W N P G S A - R T I N Y G G S F N P S G N G Y L A V Y G	AfuXyn2
83	K G W N P G S D - H D I T F S G S F N P S G N A Y L S V Y G	AfuXyn5
81	K G W Q P G T K N K V I N F S G S Y N P N G N S Y L S V Y G	Xyn2
	<u>N</u>	
117	W T T N P L I <u>E</u> Y Y V V <u>E</u> S Y G T Y N P G S G G T F R G T V	AfuXyn2
112	W T T N P L V <u>E</u> Y Y I L <u>E</u> N Y G S Y N P G S G M T H K G T V	AfuXyn5
111	W S R N P L I <u>E</u> Y Y I V <u>E</u> N F G T Y N P S T G A T K L G E V	Xyn2
147	N T D G G T Y N I Y T A V R Y N A P S I E G T K T F T Q Y W	AfuXyn2
142	T S D G S T Y D I Y E H Q Q V N Q P S I V G T A T F N Q Y W	AfuXyn5
141	T S D G S V Y D I Y R T Q R V N Q P S I I G T A T F Y Q Y W	Xyn2
177	S V R T S K R T G G T V T M A N H F N A W S R L G M N L G T	AfuXyn2
172	S I R Q N K R S S G T V T T A N H F K A W A S L G M N L G T	AfuXyn5
171	S V R R N H R S S G S V N T A N H F N A W A Q Q G L T L G T	Xyn2
	<u>A</u>	
207	H N Y Q I V A T <u>E</u> G Y Q S S G S A S I T V Y	AfuXyn2
202	H N Y Q I V S T <u>E</u> G Y E S S G T S T I T V S S G G S S S G G	AfuXyn5
201	M D Y Q I V A V <u>E</u> G Y F S S G S A S I T V S	Xyn2
228		AfuXyn2
232	S G G S S S T T S S G S S P T G G S G S C S A L W G Q C G G	AfuXyn5
222		Xyn2
228		AfuXyn2
262	I G - W S G P T C C S S G T C Q V S N S Y Y S Q C L	AfuXyn5
222		Xyn2

FIG. 55B

Day 3 Glucose Oligomers and Monomers

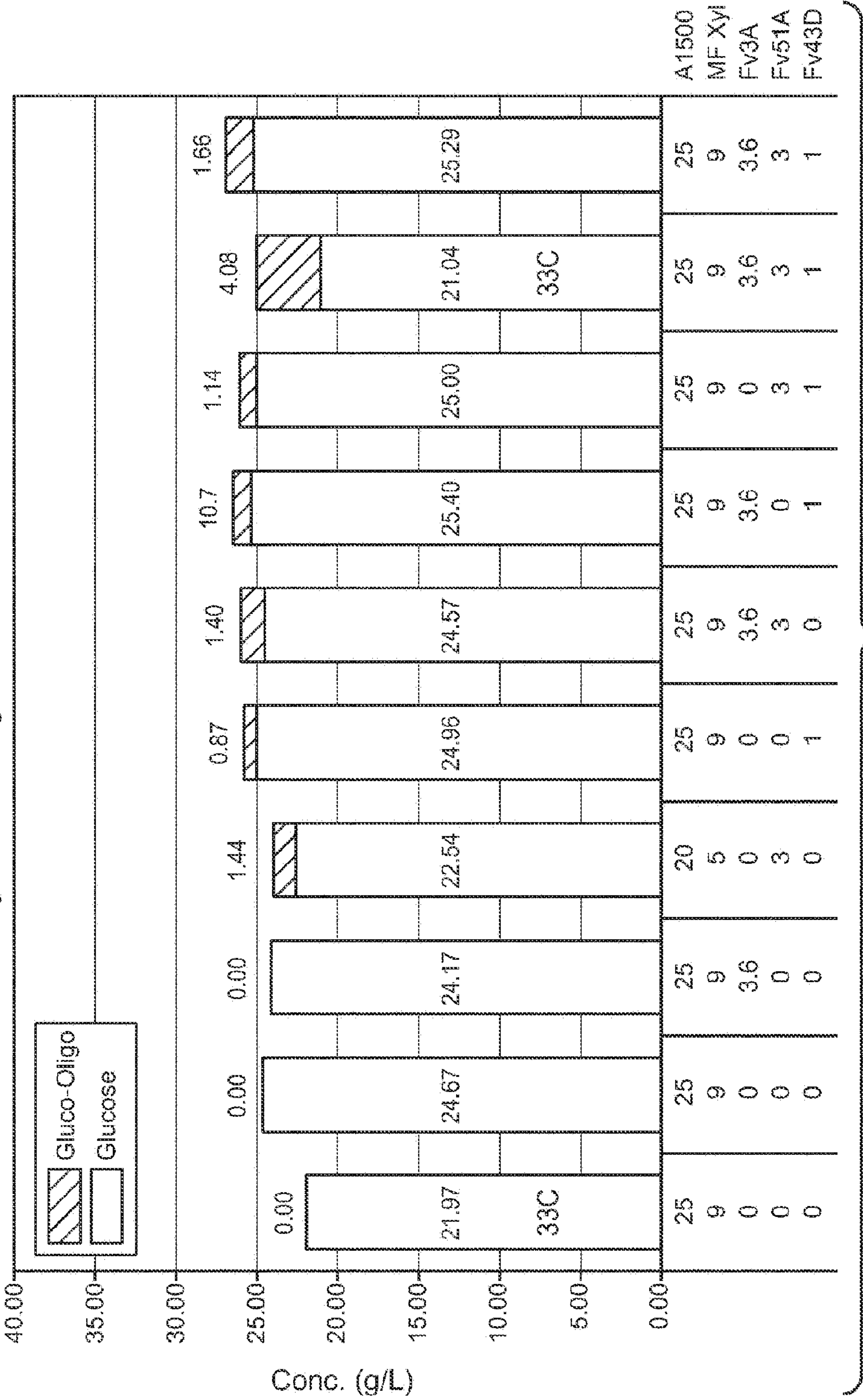


FIG. 56A

Day 3 Xylose Oligomers and Monomers

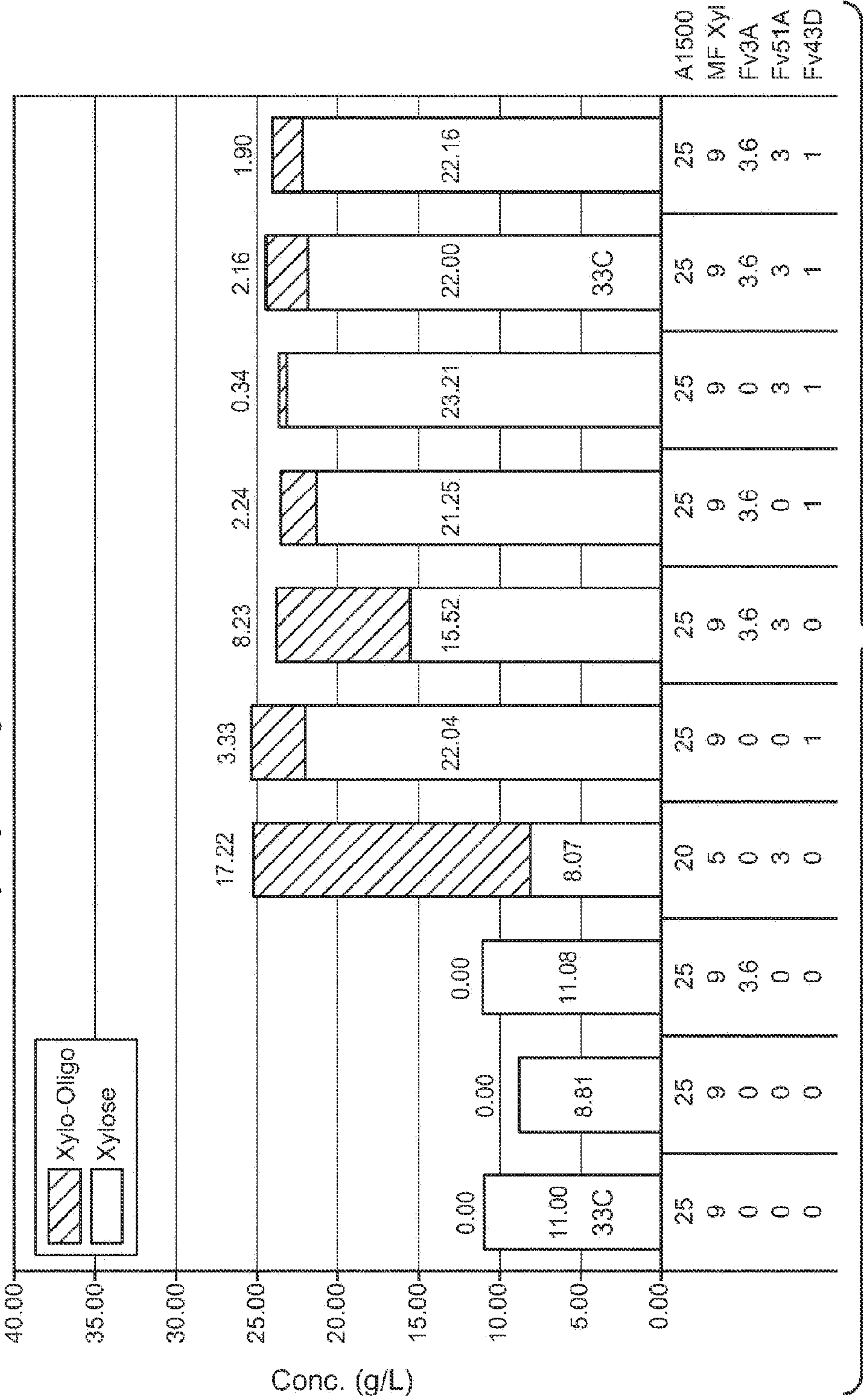
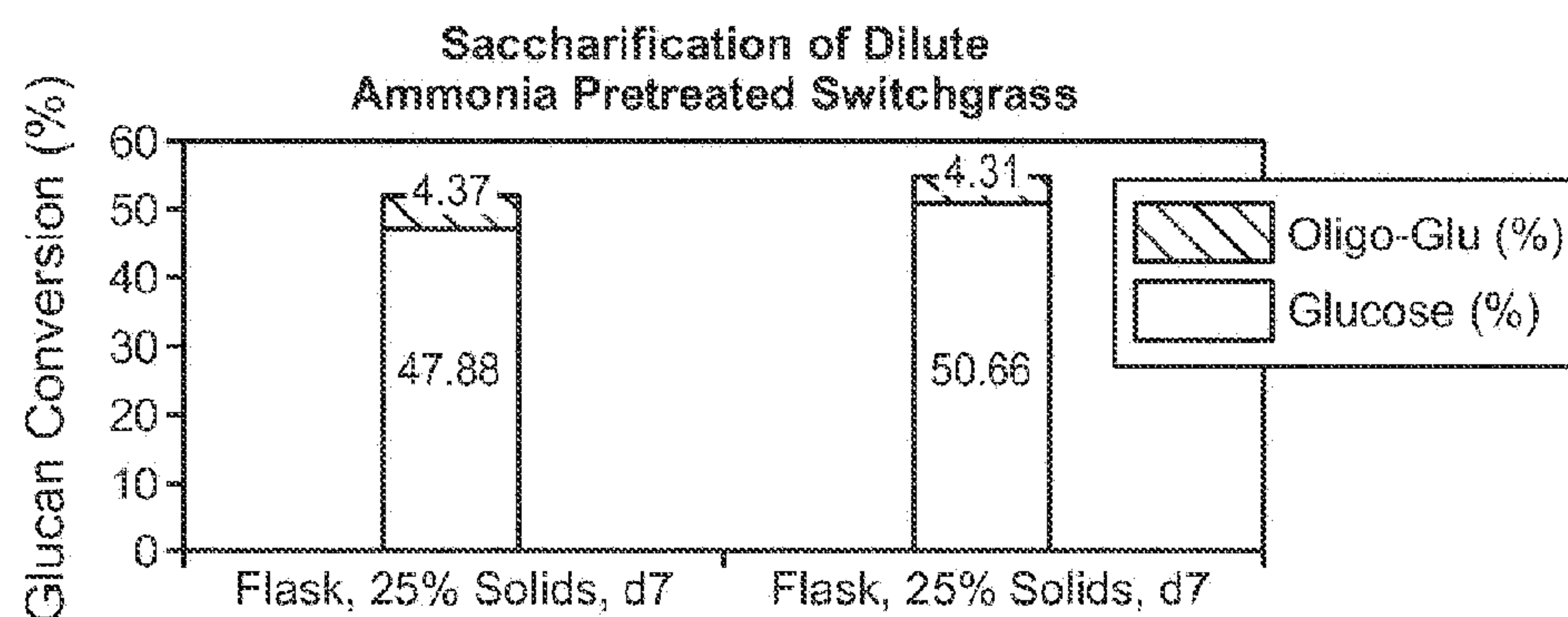
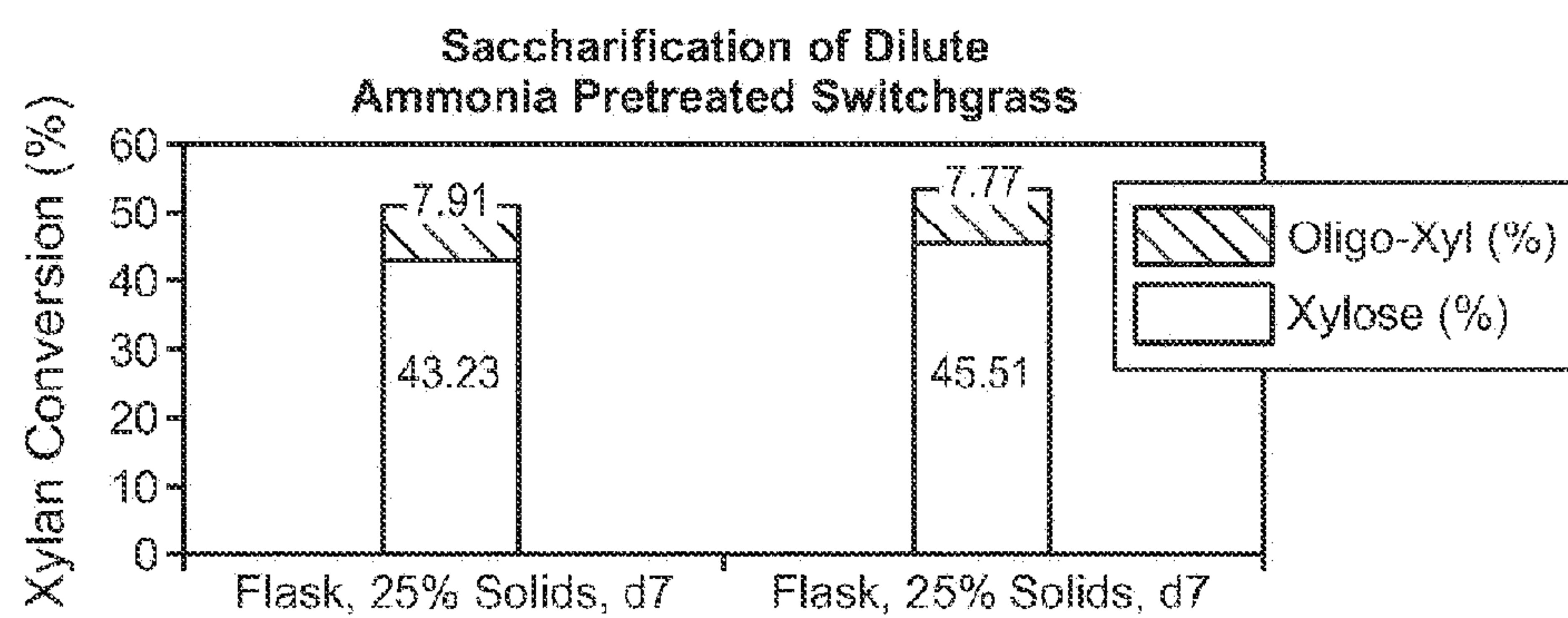
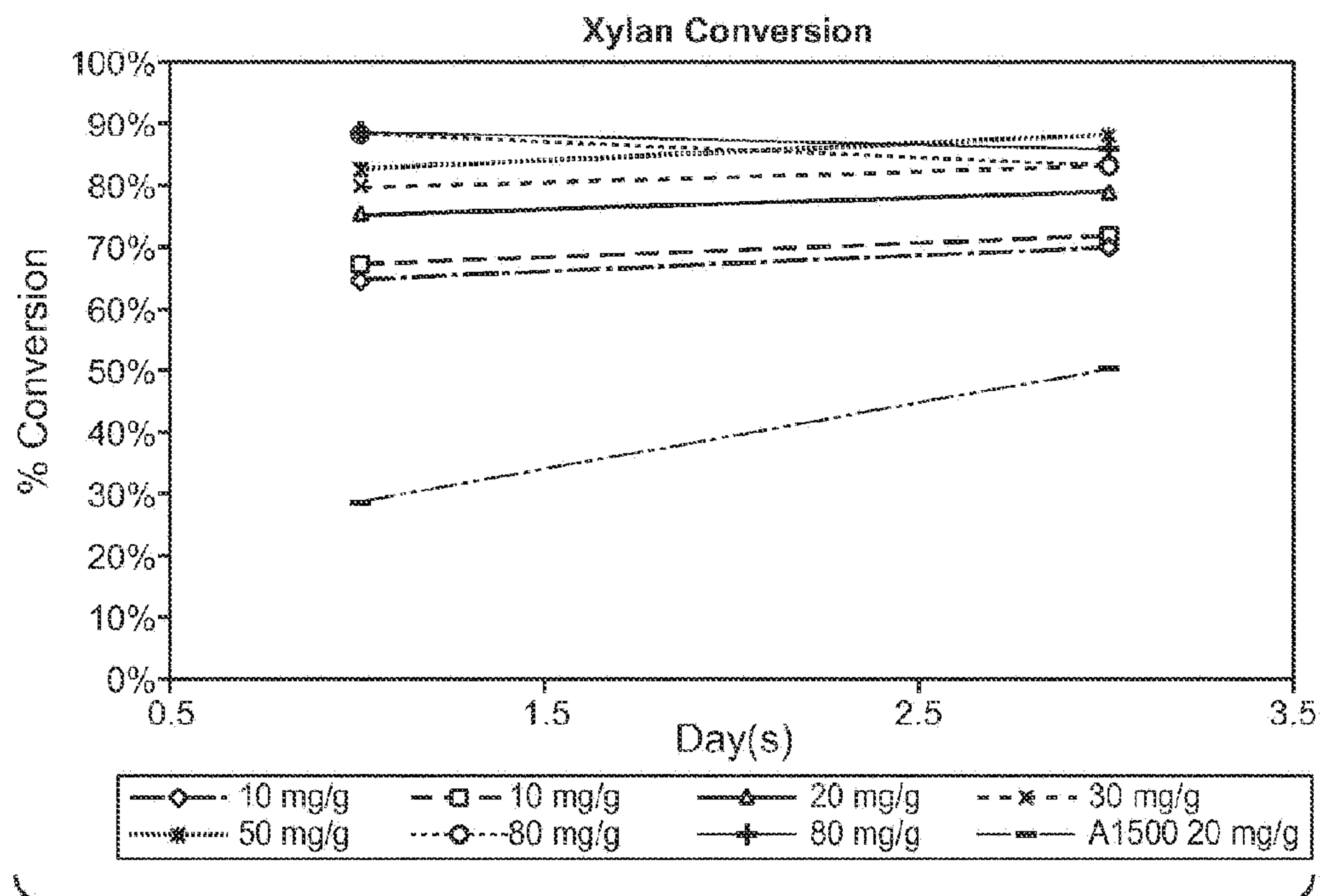
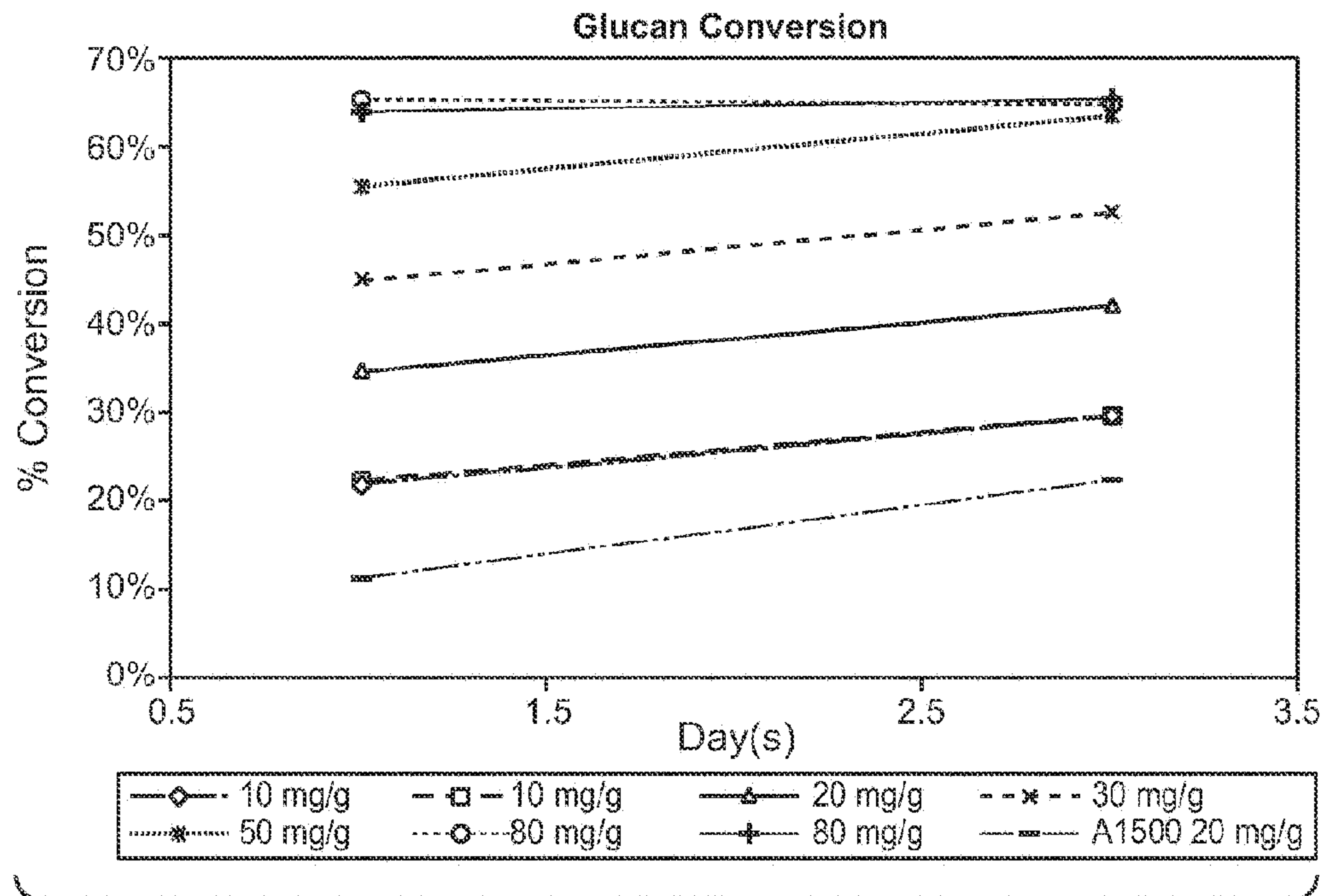


FIG. 56B

**FIG. 57A****FIG. 57B**



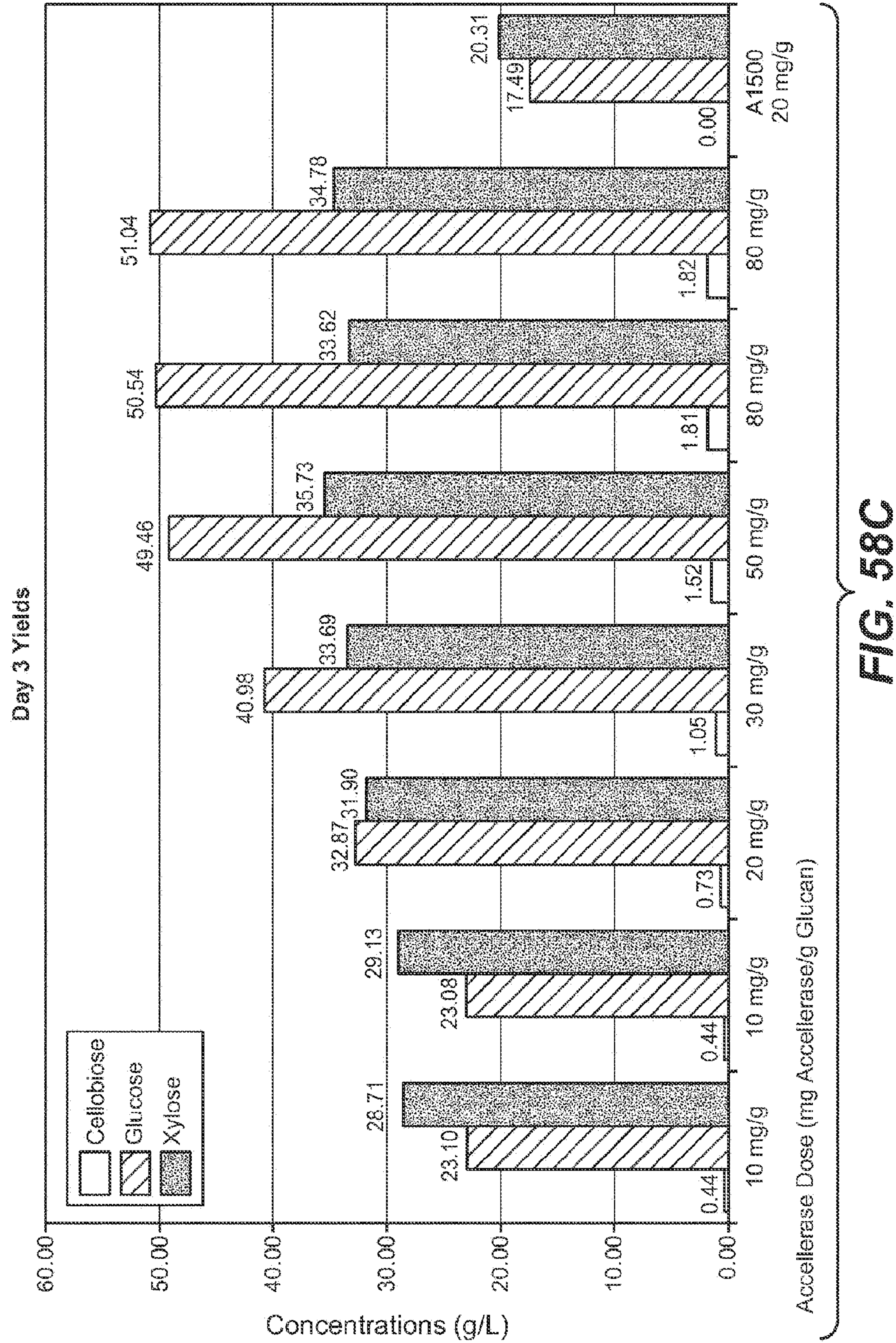
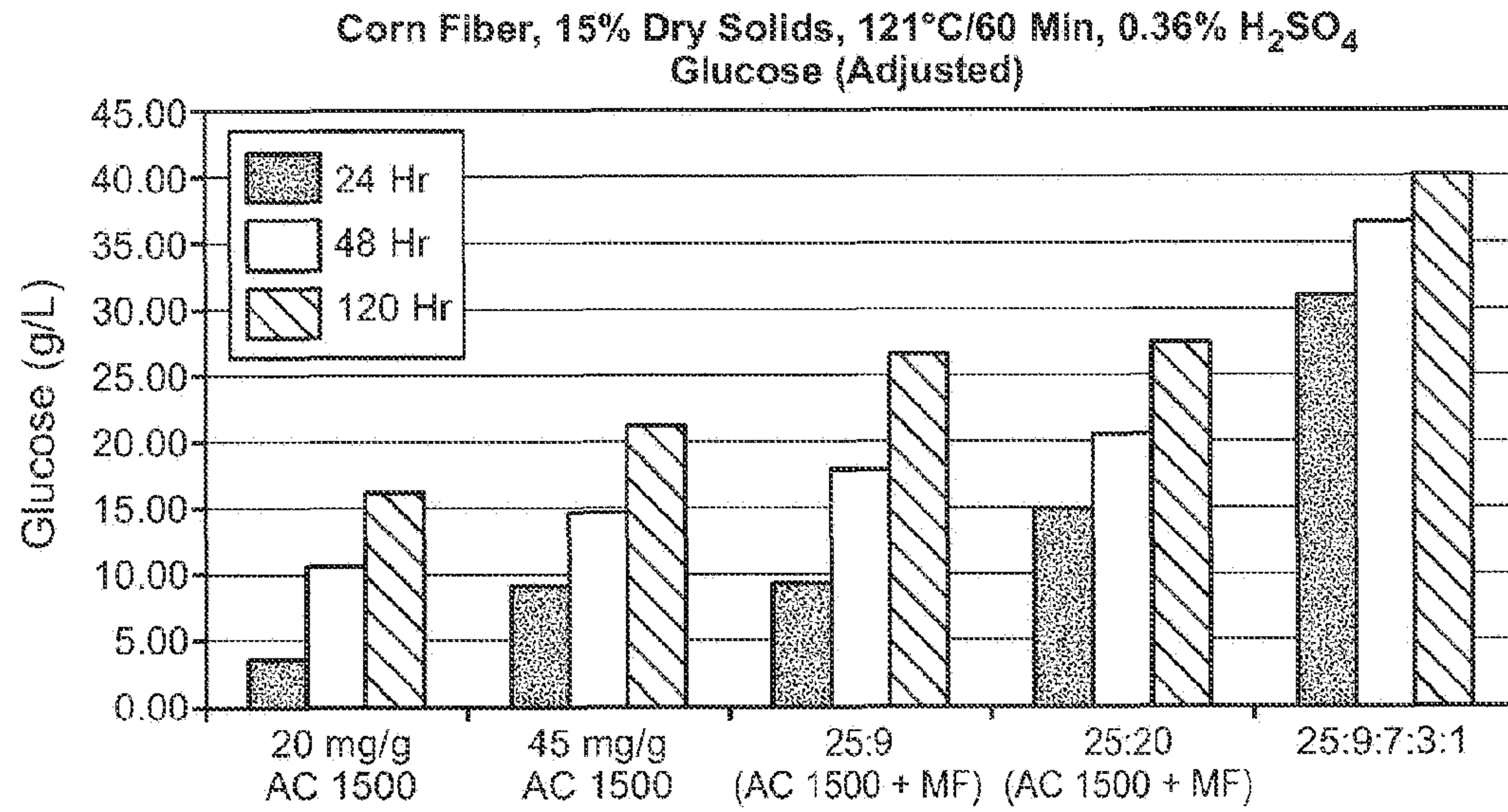
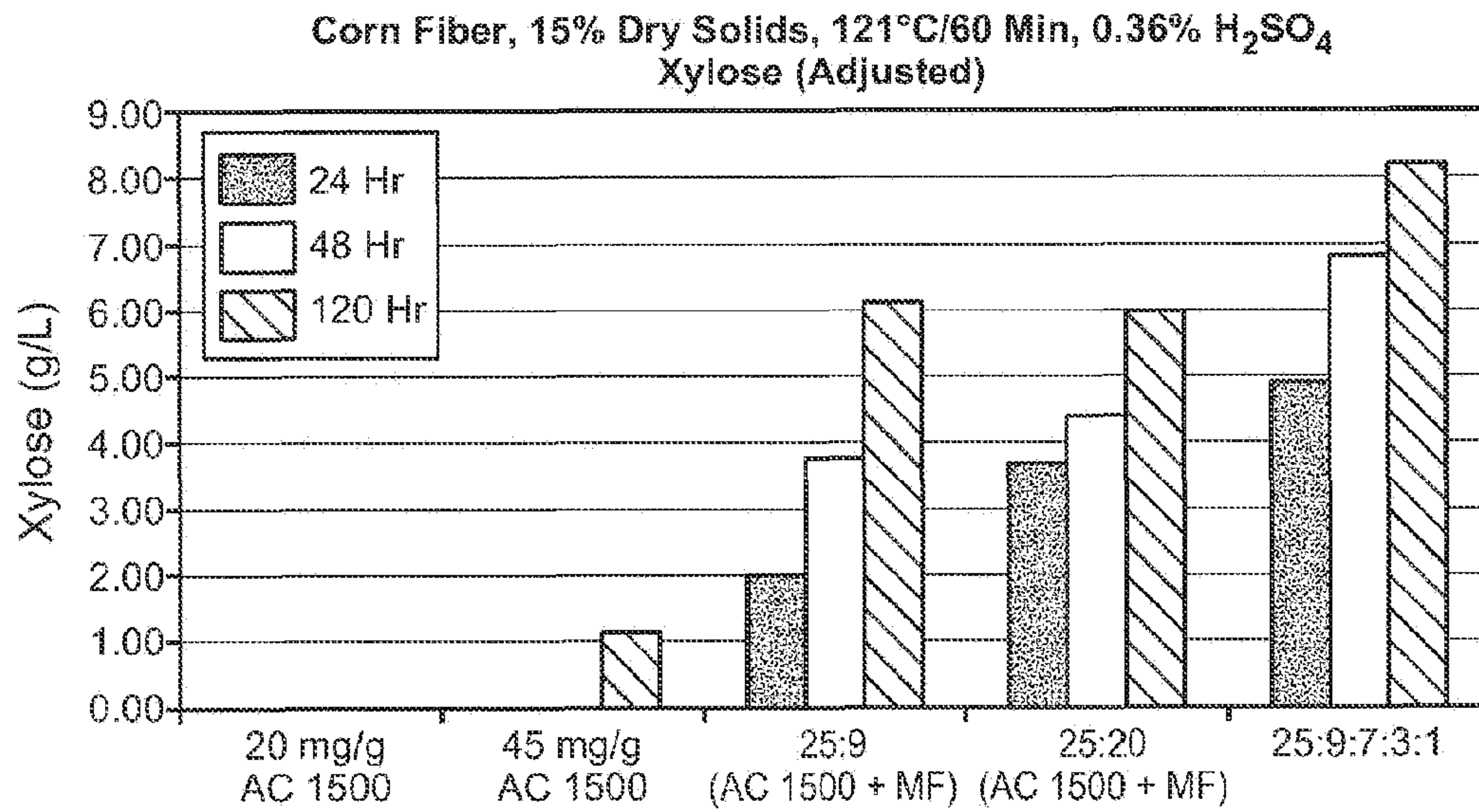
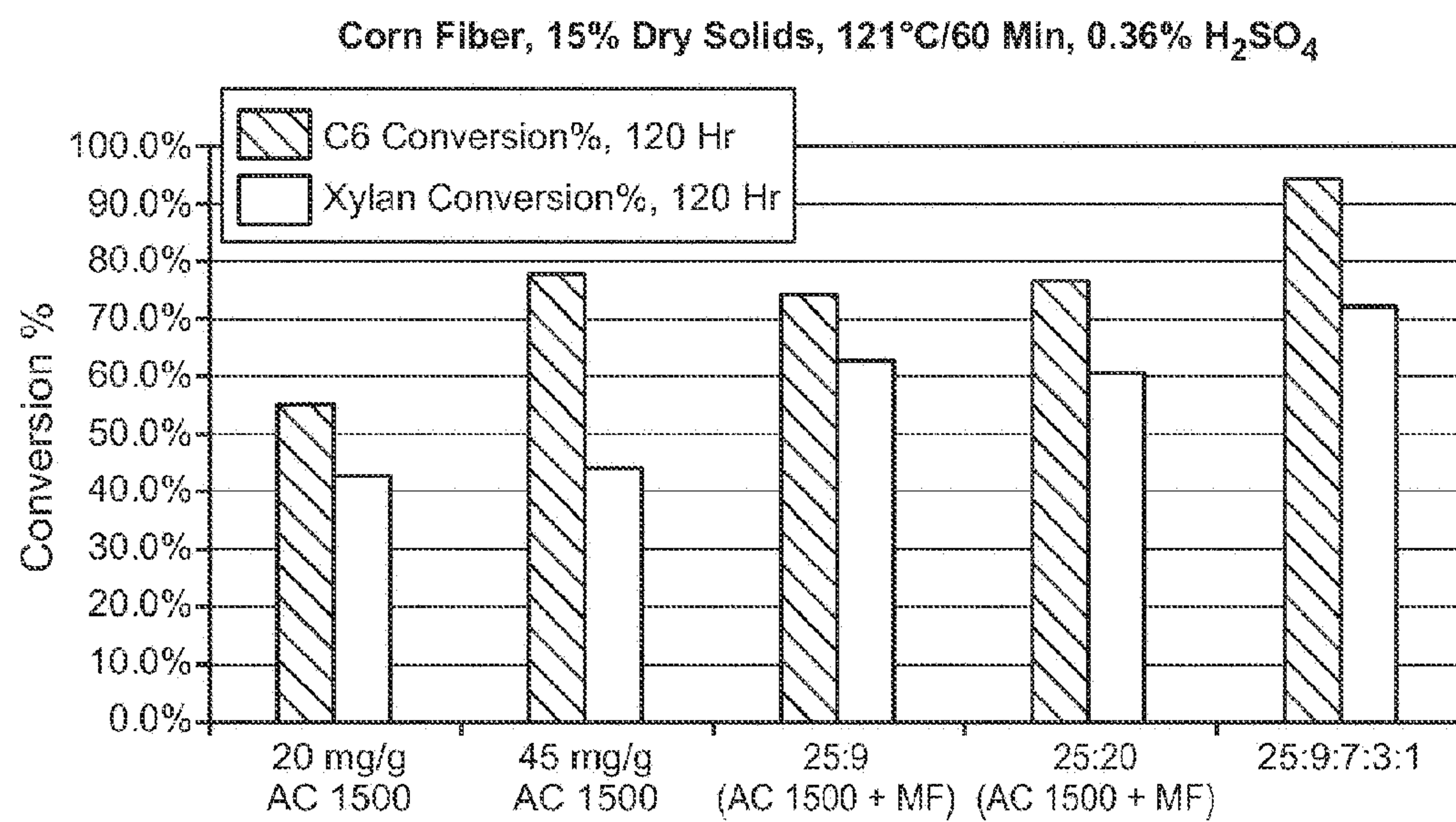


FIG. 58C

**FIG. 59A****FIG. 59B**

**FIG. 59C**

SEQ ID NO:46
Nucleotide sequence for Pa51A, a GH51 family enzyme from *Podospora anserina*

atgatccaactcaagccaagccctcgcggcgttggtggcgctgtcgacgcaatgtgtggctattgatgttggccaagtcttcggg
ggggaataagacgactgatatcatgtatggtcttattgcacgaggatatcaacaaactcggcgacggggcatctacgcgagctaa
tctccaacccgcgcgttccaagggagtgagaagttcccctccaaactcgacaactggagcccctcggtggcgctacccttaccctt
cagaagcttgccaagccccttccctctgcgttgccttactcctcaatgttgccaaaccccaaggagggcaagggaaggga
caccaaaggggaaggttggttggccaatgctgggttttgggggtatggatgtcaagaggcaagatacaactggttagcttccacg
ttactggtgagtacaagggtgactttgaggttagcttgcgcagcgcgattaccgggggagacctttggccaagaaagggtgaagggt
gggagtaagaaaggggaagtggacccgagaaggagtttgagttggtgcctttcaaggatgcgcccaaacgaacacaccttbtgtgt
gcagtgggatgccgagggcgcaaaaggacggatctttggatatctcaacttgatcagcttgttccctccgacattcaagggaaggaa
atgggctgagaattgatcttgcgcagacgatgggttgagctcaagccgaccttcttgcgcttcccgggtggcaacatgctcgaagggt
aacaccttggacacttgggtggaagtgttaacgagaccattggccctctgaaggatcgcccgggcgaatggctggtgtctgggaggtacca
gcaaaccttggcttgggtctggtcgagtaacatggagtgggccgatgacatgaacttggagcccattgtcgggtgtcttcgctggtc
ttgcctcgatggctcgttcgttcccgaatccgagatgggatgggtcatccaacaggtctctcgacgaaatcgagttcctcaactggc
gatgctaagaccaccaaatggggtgcggtcgcgcgaagcttggtcaaccccaagccttggaaggtcaagtgggttgagatcgggtaa
cgaggattggcttgccggacgcctgctggttcgagtcgtacatcaactaccgcttccccatgatgatgaaggccttcaacgaaa
agtaaccccgacatcaagatcatcgctcgccctccatcttgcacaaacatgacaaatcccgcgggtgctgcccgtgatcaacacccg
tacctgactcccgatgagttcgttgagcgaattcgccaaagttcgataacttgagcaaggataacgctgacgctcatcggcgaggctgc
gtcgacgcatoctaacgggtggttatcgcttgggaggagatatctccttgcccttgggtggcggcagtggttgcctgaggcttatct
tcttgatcagcactgagagaaacgggtgacaagatcatcggtgctacttacgcgcctggtcttgcgagcttggacccgtggcaatgg
agcatgacctgggtgcagcatgcgcgcgaccccgccctcaccactcgctcgaccagttgggtatgtctggagaaatcctcgcccacca
catcatccgtgagacgctcccgttcgatgccccggccggcacaagcccaactttgacctctgttctacgcttgccggaaagagcgaga
gtggcacccggtatcttcaaggctgcggtctacaactcgactgaatcgatcccgggtgctgttgaagtttgatggtctcaacgagggga
gcggttgcacacttgacggtgcttactgggcgggaggatccgttatggatacaaacgaccccttcaactggttatcaatggttgc
gaagaccaccttcatcaaggccgggaaaggcggaagttcaccttcaacctgcccgggttgagtggttgctgtgttgagagacggccg
acgcggtcaagggtggcaaggggaaagggcaaggggcaaggggtaactga

FIG. 60A

SEQ ID NO:47

Codon optimized cDNA for Pa51A, a GH51 family enzyme from *Podospira anserina*

atgatccacctcaagcccgccctcgcgcgcctcctcgccctcagcaccacaatgcgctcggccatcgacctct
 tcgtcaagagcagcggcggaacaagaaccacgacatcatgtacggcctcatgcacgaggacatcaacaa
 cagcggcgagcggcgcatctacgcgagctgatcagcaacgcgcgccttcacagggcagcgagaagttcccc
 agcaacctcgacaactgggtcccccgtcggcggcgcaccctcaccctccagaagctcgccaagccccctgt
 cctctgcctcccccactccgtcaacgctcgccaaccccccaaggagggtaagggtaagggcaaggacaccaa
 gggcaagaaggctggcctcggccaacgcgcgctttttggggcatggacgtcaagcggccagaaatacaccggc
 agcttccacgtcaacggcgagtagaaggcggaacttcgaggtcagcctccgcagcgccattaccggcgaga
 ccttcggcaagaaggctcgtcaaggcggcagcaagaaggcgcaagtggaccgagaaggagttcgagctggt
 ccccttcaaggacgcccccaacagcaacaacaccttcgtcgtccagtgggacgcccagggcgccaaggac
 ggcagcctcgacctcaacctcatcagcctcttccgcgcaccttcaaggcgccgcaagaacggcctccgca
 tcgacctcgcccagaccatgggtcgagctgaagcccaaccttccctccgctttcccgggcgcaacatgctcga
 gggcaacacccctcgacaacctgggtggaagtggtaacgagaccatcgccccctgaaggacggccctggcatg
 gcggcgctctgggagtagcagcagcgtgggctcggcctggctcgagtagatggagtgggcccagcagaca
 tgaacctcgagcccatcgtcggcgctctttgctggcctggcctggatggcagctttgtccccgagagcga
 gatgggctgggtcatccagcaggctctcgatgagatcgagttccctcaccggcgacgccaagaccaccaag
 tggggcgccgtccgcgccaagctcgggccaacctaaagccctggaaaggtcaaatgggtcgagatccgcaacg
 aggactggctcgccggccgacctgcgcgcttcgagagctacatcaactaccgcttcccatgatgatgaa
 ggccctcaacgagaaataccccgacatcaagatcattggccagccccctccatcttcgacaacatgacctt
 ccagcgggtgctgcgggtgaccaccacccctacctaccccccgacgaattttgtcgagcgcttcgccaagt
 tcgacaacctcagcaaggacaacgtoacctcattggcgaggccgcgcagcaccaccccccaacggcggcct
 tgcctgggagggcgacctcatgccccctgcctgggtggggcggcagcgtcgccgaggccatcttctctc
 agcaccgagcgcaacggcgacaaagatcatcggcgcacctaagccccctggcctccgatctctcgacgct
 ggcagtggagcatgacctgggtccagcaagcgcgcgcacctgcctcaccacccgcagcaccagctggta
 cgtctggcgcatcctcgcccaaccacatcattcgcgagacctcctccgtcgacgcccccgccggcaagccc
 aacttcgacccccctcttctacgtcgtggcaagtcggagagcggcacccgcatcttcaaggccgcgctct
 acaacagcaccgagagcctccccgtcagcctcaagttcgacggcctcaacgagggcgccgtcgccaacct
 caccgtcctcaccggcccccgaggaccttaacggctacaacgaccttccaccggcatcaacgtcgtcaag
 gaaaagaccaccttcatcaaggccggcaaggcgcggaagttacctttacctccccggcctctctgtcg
 ccgtcctcgagaccgcgcgacgcgctgaagggtggcaagggaagggaagggaagggaagggttaagggtacta
 a

FIG. 60B

SEQ ID NO:48**Nucleotide sequence for Gz43A, a GH43 family enzyme from *Gibberella zeae***

atgtatcggaaqttggccgcatctcggccttcttggccacagctcgtgctaccaacgacgaactgtcctc
 tcatcactagtagatggactgcgcatccttcggctcatgtctttaacgacaccttgtggctctaccgctc
 tcatgacatcgatgctggatttgagaatgatcctgatggaggccagtaacgccatgagagattaccatgtc
 tactctatcgacaagatctaagggttcctgocggctgatcacggtacggccctgtcagtgaggatgtcc
 cctgggcctctcgacagatgtgggctcctgaacgtgcccacaagaacggcaaataactaactatacttccc
 tgccaaagacaaggatgatattcttcagaatcgggcgttgctgtctcaccaacccccggcggaccattcgtc
 ccggacaagagttggatccctcacacttccagcatcgaccccgccagtttctgtcgatgatgatgacagag
 cctacttggcatggggtggatcatgggtggccagcttcaacgatggcaggataagaacaagtacaacga
 atctggcactgagccaggaaacggccacgcctgccttgagccctcagattgccaaagctgagcaaggacatg
 cacactctggcagagaagcctcgcgacatgctcattcttgaccccaagactggcaagccgctcctttctg
 aggatgaagaccgacgcttcttcgaaggaccctggattcacaaagcgcaacaagatttactacctcaccta
 ctctactggcacaacccactatcttcttatgogacttcaaagacccccctatggctccttacacctaccag
 ggcagaattctggagccagttgatggctggactactcactctagtatcgtaagtaaccaggggtcagtggt
 ggtattttatcacgatgccaaagacatctggcaaggactatcttcgccaggtaaaaggctaagaagatttg
 gtacgatagcaaaaggaaagatcttgacaaagaagccttga

FIG. 61**SEQ ID NO:49****Nucleotide sequence for Fo43A, a GH43 family enzyme from *Fusarium oxysporum***

atgtatcggaaqttggccgcatctcggccttcttggccacagctcgtgctcaagacactaatgacatto
 ctcccctgateaccgacctctgtgctcgcagatccctcggctcatgttttcgaaggcaagctctgggttta
 cccatctcacgacatcgaaagccaatgttgtaacggcacaggaggcgtcaatacgccatgagggattac
 catacctactccatgaagagcatctatggtaaagatcccggttgctgacccacggcgtcgtctctcagtcg
 atgacgttccctgggcgaagcagcaaatgtgggctcctgaacgcagctcataagaacggcaaatattatct
 gtacttccccgccaaaggacaaggatgagatcttcagaattggagttgctgtctccaacaagcccagcgggt
 cctttcaaggccgacaagagctggatccctggcaogtacagtatcgatcctgctagctacgtcgacactg
 ataacgagggcctacctcatctggggcggtatctggggcgggccagctccaagcctggcaggataaaaagaa
 ctttaacgagtcgtggattggagacaaggctgctcctaacggccaccaatgccttatctcctcagatcgcc
 aagctaagcaaggacatgcacaagatcacggaaacacccccgcgatctcgtcattctcgcccccgagacag
 gcaagcctcttcaggctgaggacaacaagcgacgattcttcgagggcccttggatccacaagcgcgga
 gctttactacctcatgtactccaccggtgataccacttccctgtctacgctacttccaagaacatctac
 ggtccttatacctaccggggcaagattcttgatcctgttgatgggtggactactcatggaagtattgttg
 agtataagggacagtggtggcttttcttctgtgatgogcatacgtctggtaaggattaccttcgacaggt
 gaaggcgaggaagatctggatgacaagaacggcaagatcttgcttcaccgctccttag

FIG. 62

SEQ ID NO:50

Nucleotide sequence for Pf51A, a GH51 family enzyme from *Penicillium funiculosum*

atgtaccggaagctcggcgtgatcagcccttctctggcactgctcgcgccatcaccatcaacgtoagcc
 agagcggcggcaacaagaccagcccgctccagtaaggcctcatgttcgaggacatcaaccacggcggga
 cggcggcctctacgcgagctgggtccggaaccgggccttccagggcagcaccgtctacccggccaaacctc
 gacggctacgactcgggtgaacggcggcattctcgcgctccagaacctcaaccaaccgctcagcccgagca
 tggcctcgtcgtcgaacgctcgccaagggctcgaaacaacggcagcaccgttcgccaacgaggggtgggtg
 gggcatcgagggtcaagcgcgagcgggtacgcgcggcagcttctacgtccagggcgactaccagggcgacttc
 gacatcagcctccagagcaagctcaaccaggaggtcttcgcgacggcgaagggtccgggtcgagcggcaagc
 acgaggactgggtccagtaacaagtaacgagctgggtccgaagaaggccgcagcaaacaccaaacacacct
 caccatcaccttcgacagcaagggcctcaaggacggcagcctcaacttcaacctcatcagcctcttcccg
 ccgacctacaacaacgggcggaacgggcctccggatcgacctcgtcagggccatggcggagctggaggga
 agttcctccgcttcccgggcggctcggacgtggaggggcgtccaggcccgtaactgggtacaagtggaaaga
 gaccgtcggcgacctcaaggaccgctactcgcgcgcgagcgcctggacctacgaggagagcaacggcacc
 ggctcctcaggtacatgaactgggtcgcgacgacatgggcctcagagccgatactcgcgcgtctgggaacggcc
 actacctcagcaacgaggtcatcagcgagaacgacctccagccgtacatcgacgacacctcaaccagct
 cgagttcctcatgggcgcgcgcggacactccctacgggtcttggagggttagcctcgggtaccggaagccg
 tggaccatcaactacgtcagagatcggaacgaggacaacctctacggcggcctcagagacctacatcgct
 accggttccaggcctactacgaacgcctacaccgccaagtaccgcacatgaccgtcatggagagcctcac
 cgagatgcgcggcccccgtgcgcggcgtcggactaccaccagtaactcgacgcgcgacggcttcgtcagc
 cagttcaactacttcgaccagatgcgcggtcaccaacgcgcagctgaacggcgagatcgccaccgtctacc
 ccaacaacccgagcaactcgggtggcgtggggcagcccggttcccgctctacccggtgggtggatcgggtccgt
 ggctgaggccgtcttctcctcagggcgaggagcggaaacagcccggaagatcatcggcgcagctacgcccc
 atgttcgcgaacattaacaactggcagtgagcccgacacctgatcgcttcgacgcgcgacagcagccgga
 cgtcgcgctctacttctggcagctcatcaagctcctcagcaccacaagatcaccacagaacctgcccac
 gacgtgggtctgggggggacatcgcccgctctactgggtcgcgcggcgggaacgacaacacccggcagcaac
 atcttcaaggccgcgcgtctacaacagcaccagcagctcccggtcaccgtccagttcgcgcggctgcaacg
 ccaagagcgcgaacctcaccatcctctcgtcggacgaccccaacgcgcagcaactacccgggcggcccccga
 ggtcgtcaagaccgagatccagagcgtcaccgccaacgcgcacggcgccttcgagttcagcctcccgaaac
 ctgtcgggtggctgtgtgaagacggagtag

FIG. 63

Fv3A MLLNLQVAASALSLSLLGGLAEAAATPYT-----LPDCTKGPLSK
Bx11 -MVNNAALLAALSALLPTALAQNNQTYANYSAQGQPDLYPETLATLTLSFPDCEHGPLKN
Bgl1 MRY-RTAAALALATG-----PFARADSHSTSG---ASAEAVVP-----

Fv3A NGICDTSLSPAKRAAALVAALTPEEKVGNLVSNAATGAPRIGLPRYNWWNEALHGLAGSPG
Bx11 NLVCDSSAGYVERAQALISLFTLEELIINTQNSGPGVPERLGLPNYQVWNEALHGLDRAN-
Bgl1 -PAGTPWGTAYDKAKAALAKLNLQDKVGIVSGVGWNGGPCVNTSPASKISYPSLCLQDG

Fv3A GRFADTP-PYDAATSEFMPILLMAAAFDDDLIHDIGNVVGTEARAFTNGGWRGVDFWTPNV
Bx11 --FATKGGQFEWATSFFMPILTAAALNRTLIHQIADIISTQARAFSNSGRYGLDVYAPNV
Bgl1 ---PLGVRYSTGSTAFTPGVQAASTWDVNLRERGGQFIGEEVKASGIHVILGPVAGPLGK

Fv3A NPFKDPRWGRGSETPGEDALHVS-RYARYIVRGLEG--DKEQRRIVATCKHYAGNDFEDW
Bx11 NGFRSPLWGRGQETPGEDAFFLSSAYTYEYITGIQGGVDPEHLKVAATVKHFAGYDLENW
Bgl1 TPQGGRNWEGF---GVDP--YLTGIAMGQTINGIQS-----VGVQATAKHYILNEQELN

Fv3A GGFTR-----HDFDAKITPQDLAEYYVRPFQECTRDAKVGSIMCAYNAVNGIPACA
Bx11 NNQSR-----LGFDAIITQQDLSEYYTPQFLAAARYAKSRSLMCAYNVNGVPSCA
Bgl1 R-----ETISSNPDDRTLHELYTWPFADAVQAN-VASVMCSYNKVNTTWACE

Fv3A NSYLQETILRGHWNWTRDNNWITSDCGAMQDIWQNEHKYVKTNAEGAQVAFENGMDSSCEY
Bx11 NSFFLQTLRESWGFP-EWGYVSSDCDAVYNVFNPHDYASNQSSAAASSLRAGTDIDCGQ
Bgl1 DQYTLQTVLKDQLGFP---GYVMTDWNNAQHTTVQSANSGLDMSMPG-TDFNGNNRLWGPA

Fv3A TTTSDVSDSYKQGLLTEKLMDRSLKRLFEGLVHTGFFDGAKAQWN-----
Bx11 TYPWHLNESFVAGEVSRGEIERSVTRLYANLVRLGYFD-KKNQYR-----
Bgl1 LTNAVNSNQVPTSRVDDMVTRILAAWYLTGQDQAGYPSFNIS-----

Fv3A --SLSFADVNTKEAQDLALRSAGEGAVLLKNDG-TLPLKLKKKDSVAMIGFWAN-----
Bx11 -SLGWKDVVKTDAWNISYEAAGEGIVLLKNDG-TLPLSKKVR-SIALIGFWAN-----
Bgl1 -----RNVQGNHKTNVRAIARDGIVLLKNDANILPLKKPASIAVVGSAAILGNHARNSP

Fv3A -----DTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVANGPTLQNSSSHDNWTTN
Bx11 -----ATTQMQGNYYGPAPYLISPLEAAKKAGYHVNFELGTEIAGNSTTG--FAK
Bgl1 SCNDKGCDDGALGMGWGSGAVNYPYFVAPYDAINTRASSQGT--QVTLSTNDNTS---SG

Fv3A AVAAAKKSDYILYFGGLDAS-----AAGEDRDRENLDWPESQLTLLQKLSSLGKPLVVIQL
Bx11 AIAAAKKSDAIIYLGGIDNT-----IEQEGADRTDIAWPGNQLDLIKQLSEVGKPLVVLQM
Bgl1 ASAARGKDVAIVFITADSGEGYITVEGNAGDRNNLDPWHNGNALVQAVAGANSNVIVVVH

Fv3A G-DQVDDTALLKNKKINSILWVNYPGQDGGTAVMDLLTGRKSPAGRLPVTQYP-----
Bx11 GGGQVDSSSLKSNKKVNSLVWGGYPGQSGGVALFDILSGKRAPAGRLVTTQYP-----
Bgl1 SVGAIILEQILALPQVKAVVWAGLPSQESGNALVDVLWGDVSPSGKLVYTTIAKSPNDYNT

FIG. 64A

Fv3A	-----SKYTEQIGMTDMDLRPT-KSLPGRTYRWYS-TPVLPYGFGLHYTKF
Bx11	-----AEYVHQFPQNDMNLRPDGKSNPGQTYIWYTGKPVYEFGSGLFYTTF
Bgl1	RIVSGGS-----DSFSEGLFIDYKHFDAN-----ITPRYEFgyGLSYTKF
Fv3A	QAKFKSN--KLTFDIQKLLKG--CSAQYS-----
Bx11	KETLASHPKSLKFNTSSILSAPHPGYTYS-----
Bgl1	NYSRLSVLSTAKS--CPATGAVVP-----
Fv3A	-----DTCALPPIQVSVKNTGR
Bx11	-----EQIPVFTFEANIKNSGK
Bgl1	-----GGPSDLFQNVATVTVDIANSQ
Fv3A	ITSDFVSLVFIKS-EVGPKPYPLKTLAAYGRLHDVAPSSTKDISLEWTLDNIAARRGENGD
Bx11	TESPYTAMLFVVRTSNAGPAPYPNKWLVGFDRLADIKPGHSSKLSIPIPVSALARVDSHGN
Bgl1	VTGAEVAQLYITYPSSAPR-TPPKQLRGFAKLN-LTPGQSGTATFNIRRDLSYWDTSQ
Fv3A	LVVYPGTYTLLLDDEPTQAKIQVTLTGKKAILDKWPQDPKSA-----
Bx11	RIVYPGKYELALNTDESVKLEFELVGEEVTIENWPLEEQQIKDATPDA
Bgl1	KWVVPsgsFGISVGASSRDIRLTSTLSVA-----

FIG. 64B

Fv3CB.pro MNPLSLGLAALSLLGYVGVNFVAAPFTDSNSGSEVLISVNGHVKQHQLDGFQASQAFQRAEDILCKDGLSKECTQHVLDDL 80
Fv3CA.pro ML-FSLVLPFLA-----FQASLALGDTs-----VTVDTSQKLQVIDGFGVSEAYGHAKQF---QNLGPGFPQKEGLDL 63

Fv3CB.pro LFSKDIGAGFSILRNGIGSSNSDKNFMNSIEPFSGSPGAKPHYVWDGYDSGQLTVAQEAfKRGCLKFLYGDAWSAPGYM 160
Fv3CA.pro LFNtTTGAGLSIIrNKIGCDAS-----NSITSTNTDNPDKQAVYHFDGDDGQ----- 111

Fv3CB.pro KTNHDENNGGYLCGVTGAACASGDWKQAYADYLLQWVEFYRKSGVKVTNlGFLNEPQFAAPYAGMLsNGTQAADFIrVLG 240
Fv3CA.pro ---SAQSMGRlCGTegVSCSSGDWRHRYVEMIAEYLSYKQAGIPVSHVGFlnEGD-GSDF--MLSTAEQAADVIPLlH 184

Fv3CB.pro KTIrKRGIHDLTIACCDGEGWDLQ---EDMMAGLTAGPDPAInYLSVVTGhGYVSPpNHPLSTTKKTWLTEWADLTGQFT 317
Fv3CA.pro SALQSKGLGDIKMTCCDNIGWKSQMDYTAKLAELEVE---KYLsVITsHEYSSSPNQPMNTTLfTWMSEGAANDQAFa 259

Fv3CB.pro PYTEFYNNsGQGEgMTWAGRIQTALVDANVSgFLYWiGAENSTNS--ALINMIGDKVIpSKRfWAFASfSRfARFGARRI 395
Fv3CA.pro T-AMyVNGGSNEGFTWAVKIAQGIVNADLSAYIYwEGVETNNKGSLSHVIdTDGtKfTISSILWAIAHWSRHIRFGAHRL 338

Fv3CB.pro EATSSVPLVTvSSFLNTDGTvATQVLNNDtVAHSVQLVVSgtG----- 438
Fv3CA.pro STSGVVQDTlVGAFENVDGSVVMVLtNSGTAAQTVDlGVSGSSfSTAQAFTSDAEaQMVDTKVTLSDGRVKVtVPVHGvV 418

Fv3CB.pro -----RNPHSLKPFLLDnsNDLTALKHLKATGKGSEfQTtIPPRSLVSEfVt----- 483
Fv3CA.pro TVKLTtAKSSKfVStAVSAQSApTPTSVKHTLTHQKTSStLTStAKAPTSTQTTSVVESAKAVKYPPfVASKGSSKSAP 498

Fv3CB.pro -----DF 485
Fv3CA.pro KKGtKKtTTKKGSHQSHKAHSAThRRCRHGsyRRGHCTN 537

FIG. 65